Social Image Analysis from a Non-IID Perspective

Zhe Xu, Ya Zhang, Longbing Cao, *Member, IEEE*

Abstract—An image in social media, termed *social image*, shows characteristics different from those images widely discussed in image processing, which can be described by both content and social use-related attributes, called *social image attributes*, including visual contents, users, tags and timestamps. There are strong coupling relationships between social image attributes, which makes social images not independent and identical distributed (non-IID). By analyzing the relationships among these attributes, we can better understand the semantic activities conducted on such non-IID social images, hence enabling new applications including content organization, recommendation and social activity understanding. In this article, we present a novel algorithm to analyze the coupling relationships between social images based on coupled similarity metrics. It involves not only intra-coupled similarity within a social image attribute, but also inter-coupled similarity between attributes, in analyzing the non-iidness of the similarity between social images. In particular, we propose a multi-entry version of coupled similarity metric to deal with attributes (i.e. tags) which have many-to-one relationship with respect to images. Experimental results on the Flickr group dataset show that the proposed algorithm achieves promising results in various applications including image clustering and tagging.

Keywords—similarity metric, social media, non-IIDness learning, structure mining

I. INTRODUCTION

THE Internet is in an era of public sharing and social interacting, encouraging more and more people to upload images or videos to media-sharing sites like Flickr and YouTube. As a result, the volume of community-contributed multimedia resources available on the web has been drastically increased. These collections raise new opportunities as well as new challenges to multimedia research.

In multimedia sharing websites, social media images (social images for short) are often associated with a set of contextual metadata in addition to their visual features. While including the metadata is expected to improve the accuracy for social image analysis, these images shall no longer be treated as independent and identical distributed (IID). For example, images posed by a user in a certain time frame may be of the same theme and hence share similar tags. When measuring the similarity of social images, one needs to consider the coupling relationship among individual attributes as well as individual images. The inter-relationships between attributes form a huge relationship network.

Zhe Xu and Ya Zhang are with Institute of Image Communication and Network Engineering and Shanghai Key Laboratory of Multimedia Processing and Transmissions, Shanghai Jiao Tong University, China, e-mail: {xz3030,ya_zhang}@sjtu.edu.cn

Longbing Cao is with the Advanced Analytics Institute, University of Technology, Sydney.

Several studies have previously attempted to analyze social images utilizing their contextual metadata. The relationship between visual features and tags are mined to annotate web images [1] and measuring the distance between semantic concepts [2]. Wu et al. cluster web images by mining correlations between images and their surrounding texts [3]. Recently, multi-relational clustering techniques have been proposed which consider multiple objects in heterogeneous networks [3], [4]. Most of the existing methods dwell on modeling a unidirectional link between two attributes, such as how visual similarity influences image tags. Fig. 1 shows typical strategies dealing with non-IID social images.

There are strong non-IID characteristics embedded in social image applications [5]. None of existing work fully captures coupling relationships between attributes, values and images. There is a lack of methods that can utilize and analyze various kinds of metadata in a unified view. A fundamental issue shared by social image applications is how to measure the similarity between attributes, values and images in social image applications. We try to find a way to model the whole relationship network from different perspectives. Similarity metric is a fundamental issue in image analysis. For instance, content organization can be formulated as an image clustering problem, which can be solved in a bottom-up manner using image similarity. Network analysis contains the generation of tag network and user network, each of which is an ensemble of links between similar tags or users weighted by a pairwise distance. In addition, the discovery process can be considered as a multidimensional recommendation employing similarity on each dimension. We thus propose a coupled similarity metric for social images from a non-IID perspective to support various social image applications.

For non-IID applications, instead of managing each attribute individually, the coupled object similarity (COS) [6] enables interaction and feedback from other attributes. For social image analysis, COS has an appealing feature that it not only captures the similarity among objects (images), but also provides intermediate results of similarity between attribute values (e.g. tags and users). However, COS is not directly applicable to measure the similarity among the objects for which an attribute may have multiple values (called a *multi-entry* attribute). For instance, in Flickr, the number of tags for an image can range from zero to thousands. Hence, the tag-image relationship, which plays a vital role in practice, is a many-to-one relationship. A similar situation exists when describing movies by genres [7]. The authors solve the problem by converting multi-entry attributes into several supportive Boolean attributes, where 0 for absence and 1 for presence of a value. The modification may lead to extremely high time complexity, making it inapplicable to attributes with a large set of values such as the tag attribute of a social image.
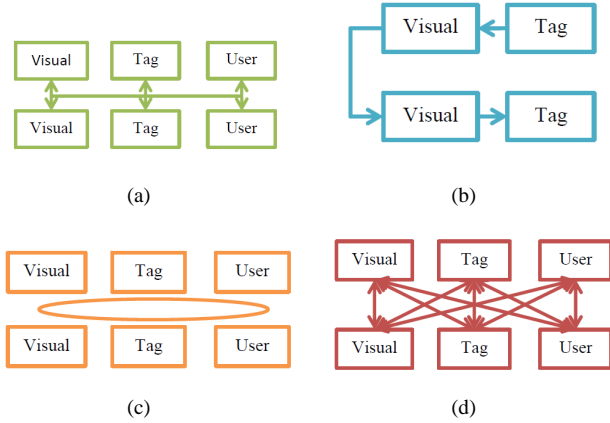
Fig. 1. Demonstration of different methods dealing with non-IID. social images. (a) Algorithms consider multiple attributes, but not introduce coupling relationship between them. (b) Propagation algorithms employ unidirectional links from one attribute to another. (c) Clustering algorithms put multiple attributes together in a black box. It is hard to provide an explicit explanation for why an algorithm succeeds or fails. (d) Coupled analysis introduces coupling relationships within multiple attributes.

We propose in this paper a coupled similarity metric for the multi-entry attributes named *Multi-entry Coupled Object Similarity* (MeCOS), which treats the tags of an image as a feature vector. MeCOS calculates intra-coupled and inter-coupled similarity between individual attribute values. By iteratively choosing the most similar value pair in two vectors, the similarity between value vectors is calculated based on singe value similarities. MeCOS shows promising results in comparison with a state-of-the-art algorithm based on nonnegative matrix factorization [4].

The main contributions is as follows. 1) We introduce the idea of coupled similarity metrics to social image analysis to handle the non-IID characteristics of social images. By measuring the similarity between images and between attribute values (tags, users), the algorithm provides a unified view to understand the strong semantic interactions embedded in social image collections that supports applications of image clustering, tag network, and tag recommendation. 2) To tackle the challenges brought by multi-entry attributes, the Multi-entry Coupled Object Similarity (MeCOS) metric is proposed to capture the many-to-one relationships between attribute values and objects.

The rest of the paper is organized as follows. Section II summarizes the existing studies. Section III reviews coupled similarity metrics and presents some open issues of the similarity metrics. The multi-entry coupled similarity metric is presented in Section IV. Section V briefly introduces the dataset for experiments together with visual features and metadata used in the analysis. Section VI presents the performance of MeCOS to different social image applications. Section VII introduces possible improvement and extension of the algorithm. Conclusions are drawn in Section VIII.

II. RELATED WORK

In this section, we briefly review related work on social image algorithms and applications and on analyzing coupled similarity to cater for non-IID characteristics.

Social media analysis. Many applications have now involving social media analysis, including event identification [8], city exploring [9], ecological phenomena discovery [10] and personalized recommendation [11], [12]. Due to the well known semantic gap [13], visual features alone are not able to represent images at the semantic level. Many methods have been proposed to employ various types of metadada available on multimedia sharing websites. Considering tag as the most important semantic context, many researchers aim to reduce the semantic gap using tags [14]–[16]. Meanwhile, geographical information is widely used for image annotation, scene summarization and event detection [17]–[19].

Network analysis. There are various networks associated with social images including user network, tag network and image network. To explore users' social networking behavior, Palla et al. [20] quantify the evolution of social communities based on temporal membership changes. Kumma et al. [21] characterize the user network into categories and indicate the action pattern of each user. Wu et al. present Flickr distance as a measurement of the relationship between semantic concepts in visual domain [2]. For the image network, the problem can be seen as structure mining. "Topic Models" [22] discover patterns by representing the underlying topics with word distributions. Recently, Latent Dirichlet allocation (LDA) [23] is widely used to model the topic clusters. Diffusion-based methods put images and metadata together into a bipartite graph [24] or hypergraph [3], and perform co-clustering algorithm to extract similar images.

Multi-relational algorithms The multi-relational clustering techniques learn structures given various entities and multiple relationships. Early algorithms [25], [26] fail to achieve reasonable results on large-scale datasets. Sun et al. [27] introduce a similarity measure called PathSim that is able to find peer objects in heterogeneous networks. Lin et al. [4] propose a nonnegative matrix factorization (NMF) algorithm to characterize relational structures of group photo streams. Image clustering and tag prediction are preformed to demonstrate the performance of the structure discovery process.

Non-IIDness learning More recently, increasing attentions have been paid to the coupling among attributes and objects in complex applications such as coupled behaviors in a community and semantic linkage between items for recommender systems [6], [28], [29]. The discussion in [5] further summarizes major theoretical challenges of complex behavior and social applications and so-called *Big Data* applications as heterogeneity and coupling, which forms the need of non-IIDness learning. In the relevant applications, different types of data sources make the problem not identically distributed, while various couplings between objects, attributes, relations make it not independent. By introducing intra-coupled and inter-coupled relationship within and between attributes, non-IIDness-oriented coupled analysis has been successfully applied in coupled clustering [6], coupled behavior analysis [29],

[30], item recommendation [7], recommender system [31], and coupled clustering ensemble [32].

Social image applications present strong non-IID characteristics as discussed in Section I. However, none of existing work systematically captures the coupling relationships between social image attributes, values and images. While non-IIDness learning and coupled analysis have been shown effective in handling many applications with non-IID features, we have not found the related work reported for social image applications to systematically capture comprehensive couplings between social image attributes, values and images. In particular, social images present the so-called multi-entry challenge to existing coupled similarity metrics. In this work, we explore the coupled similarity in social images with many-to-one couplings between an object and multiple values associated with an attribute.

III. COUPLED SIMILARITY METRICS

In this section, we briefly review the basic algorithm of coupled similarity metrics [6] and present possible modifications and extensions. The algorithm is proposed in terms of both intra-coupled and inter-coupled attribute value similarities. Here, "intra-coupled" stands for the relationship within an attribute (i.e. value frequency distribution), while "inter-coupled" means the interaction of other attributes with this attribute (i.e. feature dependency aggregation). The details about the functions and discussions can be found in [6].

A. Basic Information Functions

The similarity between two objects is built on top of the similarities within their values for all features. Several information functions are defined in formulating the relationship within and between multiple attributes [6].

- Three Set Information Functions (SIFs) are defined to reflet relationships between objects and attribute values.
- An Inter-Information Function (IIF) is defined to transform the information from one attribute to another attribute through the bridge of objects.
- Finally, the Information Conditional Probability (ICP) is calculated to represent the probability that different kinds of attribute values appear together.

B. Intra-coupled and Inter-coupled Interaction Similarity Metrics

To measure the similarity of values within an attribute, the algorithm considers the relationship between attribute value frequencies. Two values are similar if the frequencies are large and nearly the same.

In the view of inter-coupled interaction, two attribute values are closer to each other if they have more similar value subsets in other attributes in terms of co-occurrence object frequencies. Hence, the inter-coupled similarity between two attribute values is defined as the probability of co-occurrence in the value space of the other attributes.

C. Coupled Object Similarity

The Coupled Attribute Value Similarity (CAVS) between attributes x and y is introduced naturally as the product of intra-coupled and inter-coupled similarity. The sum of all CAVSs forms the Coupled Object Similarity(COS) between objects.

D. Open Issues

Although the coupled similarity algorithm achieves fairly good performance on applications including clustering and recommendation, there are still several open issues that the algorithm fails to address. First, for intra-coupled similarity, the algorithm only considers the occurrence of each value. It assumes that "the discrepancy of attribute value occurrence times reflects the value similarity in terms of frequency distribution" [33]. Apparently, the assumption, with the possible semantic information missing, may not be robust enough. For most of the attributes, such as visual contents, the semantic similarity is more reasonable than statistics. The second issue lies in the multi-entry problem. For specific attributes, such as tags of an image or genres of a movie, an object may be described by possibly more than one values. The standard coupled similarity algorithm fails to cover this scenario.

IV. MULTI-ENTRY COUPLED OBJECT SIMILARITY

As discussed before, the standard coupled similarity metrics do not address the multi-entry problem, i.e. attributes that have many-to-one relationship with objects. In a previous attempt to solve this problem, each possible value of the multi-entry attribute is re-assigned as a new attribute of Boolean values [7]. Therefore, the intra-coupled similarity between value vectors turns into the inter-coupled similarity between new attributes. As a consequence, the computational complexity increases dramatically. This solution works reasonably well or the application of movie recommendation because the multi-entry attribute involved (i.e. movie genre) has a relatively small set of values. However, when considering tag as an attribute, there can be hundreds or even thousands of independent values, making the computational complexity too high to deal with practical applications.

In this section, we present a multi-entry coupled object similarity algorithm. Given a multi-entry attribute, instead of turning all the possible values into new attributes, we represent it in the original form - feature vector. In that manner, we first calculate the coupled similarity between individual attribute values, then perform a championship list selection to turn the single value similarity into the similarity between vectors.

A. Problem Statement

The goal of coupled similarity analysis is to measure the similarity between data objects. Each object is described by several attributes. Here is a summarization of notations in the algorithm. Data objects are organized by an information table $S = < U, A, V, f >$, where $U = \{u_1, ..., u_m\}$ is a finite set of data objects; $A = \{a_1, ..., a_n\}$ is a finite set of features; $V = \bigcup_{j=1}^{n} V_j$, V_j is the set of attribute values of feature a_j;

TABLE III.   NOTATION TABLE

| Notation	Abbreviation	Explanation	Example
u_i	-	objects, indexed from 1 to m	$u_1, u_2, ..., u_6$
a_j	-	attributes, indexed from 1 to n	A, B, C
m	-	total number of objects	6
n	-	total number of attributes	3
V_j	-	set of all possible values of attribute j	$V_2 = [B_1, B_2, B_3, B_4]$
$f_j(u_i)$	-	The value for an object u_i of attribute j	$f_2(u_2) = [B_1, B_3]$
$\mathcal{F}_W(O)$	VCF	total count of attribute values in W that occurs for objects in O	$\mathcal{F}_{\{B_1, B_4\}}(\{u_1, u_2\}) = 1 + 1 = 2$
$f_j^*(\{u_{r_1}, ..., u_{r_t}\})$	SIF1	attribute values of a set of objects	$f_2^*(\{u_1, u_5\}) = \{[B_1, B_2], [B_2, B_3]\}$
$g_j(x)$	SIF2	objects whose attribute value is (contains) x for a_j	$g_2(B_3) = \{u_2, u_4, u_5\}$
$g_j^*(W)$	SIF3	objects whose attribute value is (contains) at least one element in W for a_j	$g_2(B_3, B_4) = \{u_2, u_3, u_4, u_5, u_6\}$
$\phi_{j \to k}(x)$	IIF	map from a value x of attribute j to attribute k by shared objects	$\phi_{3 \to 2}(C_2) = \{B_1, B_2, B_3, B_4\}$
$P_{k|j}(W|x)$	ICP	probability of attribute k, value W conditioned on attribute j and value x	$P_{2|3}(\{B_1\}|C_2) = \frac{1}{7}$
$\delta_j^{Ia}(x, y)$	IaAVS	Intra-coupled Value Similarity between values x and y of attribute j	$\delta_2^{Ia}(B_1, B_2)$
$\delta_{j|k}(x, y)$	IRS	Inter-coupled Relative Similarity between values x and y of attribute j by attribute k	$\delta_{3|2}(B_1, B_2)$
$\delta_j^{Ie}(x, y)$	IeAVS	Inter-coupled Value Similarity between values x and y of attribute j by all other attributes	$\delta_2^{Ie}(B_1, B_2)$
$\delta_j^{Aa}(x, y)$	CAVaS	Coupled Attribute Value Similarity between values x and y of attribute j	$\delta_2^{Aa}(B1, B2)$
$\delta^{Ae}(W_i, W_j)$	CAVeS	Coupled Attribute Vector Similarity between attribute vectors W_i and W_j of attribute k	$\delta^{Ae}([B_1, B_2], [B_1, B_3])$
$COS(u_i, u_j)$	COS	Coupled Object Similarity between objects u_i and u_j	$COS(u_1, u_4)$
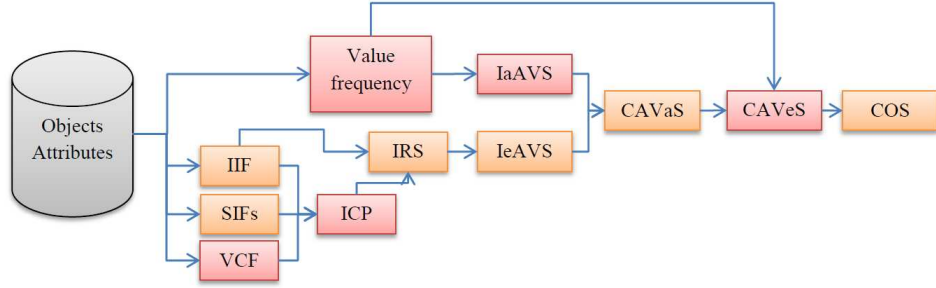
Fig. 2.   Working process diagram to describe the MeCOS algorithm. Support functions such as IIF, SIF and ICP are calculated to compute the inter-coupled similarity (IeAVS). In addition to standard COS, we introduce attribute vector similarity (CAVeS) based on attribute value similarity (CAVaS).

TABLE I.   A TOY EXAMPLE OF INFORMATION TABLE

U/A	a_1	a_2	a_3
u_1	A_1	B_1	C_1
u_2	A_2	B_1	C_1
u_3	A_2	B_2	C_2
u_4	A_3	B_3	C_2
u_5	A_4	B_3	C_2
u_6	A_4	B_2	C_3

TABLE II.   A TOY EXAMPLE OF INFORMATION TABLE WITH A MULTI-ENTRY ATTRIBUTE B.

U/A	a_1	a_2	a_3
u_1	A_1	B_1, B_2	C_1
u_2	A_2	B_1, B_3	C_1
u_3	A_2	B_2, B_4	C_2
u_4	A_3	B_1, B_2, B_3	C_2
u_5	A_4	B_2, B_3	C_2
u_6	A_4	B_1, B_2, B_4	C_3

and $f = \bigwedge_{j=1}^n f_j (f_j : U \to V_j)$ is the value of the j-th attribute of item u_i.

Let the k-th attribute a_k be an attribute with multi-entry feature. The attribute a_k has an information sub-table $S_k = <U, V_k, f_k>$. Here, since the value of a_k is a vector instead of a value, $f_k : U \to 2^{V_k}$ satisfies $\bigcup_{i=1}^m f_k(u_i) = V_k$.

Table I shows a toy example of the information table performing standard coupled similarity metrics. Table II introduces a multi-entry attribute a_2, $V_2 = \{B_1, B_2, B_3, B_4\}$ and $f_2(u_2) = \{B_1, B_3\}$.

B. Multi-entry Information Functions

For multi-entry attributes, we re-define some of the information functions. The meaning of the information functions can be found in Sec. III-A.

DEFINITION 1: Suppose the k-th attribute has multi-entry feature. Three **Set Information Functions (SIFs)** are defined as $f_k^* : 2^U \to 2^{2^{V_k}}$, $g_k : V_k \to 2^U$, and $g_k^* : 2^{V_k} \to 2^U$.

$$f_k^*(\{u_{r_1}, ..., u_{r_t}\}) = \{f_k(u_{r_1}), ..., f_k(u_{r_t})\}, \quad (1)$$

$$g_k(x) = \{u_i | x \in f_k(u_i), 1 \leq i \leq m\}, \quad (2)$$

$$g_k^*(W) = \{\bigcup u_i | f_k(u_i) \cap W \neq \emptyset, 1 \leq i \leq m\}, \quad (3)$$

where $u_i, u_{r_1}, ... u_{r_t} \in U$, and $W \subseteq V_k$.

With the respective SIFs for single-entry and multi-entry attributes, the form of Inter-Information Function(IIF) stays unchanged.

DEFINITION 2: The **Inter-Information Function(IIF)** $\phi_{j \to k} : V_j \to 2^{V_k}$ is:

$$\phi_{j \to k}(x) = f_k^*(g_j(x)), \quad (4)$$

However, when referring to the Information Conditional Probability (ICP), the original form no longer suits for multi-entry attributes. As a property of the conditional probability,

given a condition, the probability of all possible values should sum to one:

$$\sum_i P_{k|j}(w_i|x) = 1 \qquad (5)$$

where $x \in V_j$, $w_i \in V_k$ ($w_i \subseteq V_k$ in a multiple version), and $\bigcup_i w_i = V_k$.

Supposes we follow the original ICP formulation where $P_{k|j}(W|x) = \frac{|g_k^*(W) \cap g_j(x)|}{|g_j(x)|}$. Assume $G = g_j(x)$ and $G_i = g_k^*(t_i) \cap g_j(x)$. For the single-entry situation, G_i contains only one attribute value. Therefore, we have got $\biguplus_i G_i = G$ and $\sum_i P_{k|j}(w_i|x) = 1$ correctly.

In the multiple version, when attribute k has multiple entries, G_i turns to a feature vector. Since the intersection between G_i and G_j are not always empty, the sum of all conditional probabilities can exceed one.

Therefore, instead of the number of objects, we treat ICP as the proportion of the total number of attribute values in the object set.

DEFINITION 3: We define the **Value Count Function (VCF)** $\mathcal{F}_W(O) : 2^U \to \mathcal{N}$ to be the count of attribute values in $W \subseteq V_k$ which appears in the set of feature vectors from an object item subset $O \subseteq U$.

$$\mathcal{F}_W(O) = \sum_{w \in W, o \in O} \mathcal{I}_w(o) \qquad (6)$$

where $\mathcal{I}_w(o)$ is an indicator function indicates whether the attribute vector of an object o contains an attribute value w.

DEFINITION 4: Suppose that the k-th attribute has multi-entry feature. Given the k-th attribute value $W \subseteq T$ and the j-th attribute value $x \in V_j$, the **Information Conditional Probability (ICP)** of W with respect to x is:

$$P_{k|j}(W|x) = \frac{\mathcal{F}_W(g_k^*(W) \cap g_j(x))}{\mathcal{F}_T(g_j(x))} \qquad (7)$$

In Table II, $f_2^*(\{u_1, u_2\}) = \{\{B_1, B_2\}, \{B_1, B_3\}\}$, $g_2(B_1) = \{u_1, u_2\}$, while $g_2^*(\{B_1, B_2\}) = \{u_1, u_2, u_3, u_6\}$, $\phi_{3\to2}(C_2) = \{B_1, B_2, B_3, B_4\}$. If we use the original ICP, $\sum_{i=1}^4 P_{2|3}(\{B_i\}|C_2) = \frac{1}{3} + 1 + \frac{2}{3} + \frac{1}{3} = 2.33 > 1$. For the proposed ICP function, $P_{2|3}(\{B_1\}|C_2) = \frac{|\{B_1\}|}{|\{B_2, B_4, B_1, B_2, B_3, B_2, B_3\}|} = \frac{1}{7}$. Then $\sum_{i=1}^4 P_{2|3}(\{B_i\}|C_2) = \frac{1+2+3+1}{7} = 1$.

C. Intra-coupled Value Similarity

In the multi-entry situation, the intra-coupled value similarity can be formulated in a much rational way since similar values tend to appear together in an object. The most straightforward idea is the co-occurrence matrix. However, in co-occurrence matrix, the most frequently occurred tags tend to get higher co-occurrence value. In practice, these tags usually indicate abstractive concepts, it is error prone to say that they are semantic similar to other tags. We call them group related tags.

To solve this problem, we formulate the intra-coupled value similarity by introducing the idea of conditional probability.

DEFINITION 5: For a multi-entry attribute k, given the information sub-table S_k, the **Multi-entry Intra-coupled Attribute Value Similarity (MeIaAVS)** between value x and y is defined as:

$$\delta_k^{I_a}(x,y) = max(P((x,y)|x) + P((x,y)|y) - 1, 0) \qquad (8)$$

where $P((x,y)|x)$ stands for the probability that attribute value y appears in an object's attribute vector in condition that x is already in that vector.

High similarity is achieved only when two tags always appear at the same time (like "barack" and "obama"). We define the similarity value to be non-negative, $\delta_k^{I_a-sgl} \in [0,1]$.

In the toy example, B_1 occurs for 4 times, B_2 occurs for 5 times, there are 3 objects whose attribute a_2 include both B_1 and B_2. So $\delta_f^{I_a}(B_1, B_2) = \frac{3}{4} + \frac{3}{5} - 1 = 0.35$.

D. Inter-coupled Value Similarity

With the definition of IIF and ICP in Section IV-B, based on intersection set, we define the **Multi-entry Inter-coupled Relative Similarity (MeIRS)** between value x and y of an attribute j based on a multi-entry attribute k as:

$$\delta_{j|k}(x,y) = \sum_{w \in \bigcap} \min\{P_{k|j}(\{W\}|x), P_{k|j}(\{W\}|y)\}, \qquad (9)$$

where $w \in \bigcap$ denote $w \in (\bigcup \phi_{j\to k}(x)) \cap (\bigcup \phi_{j\to k}(y))$.

DEFINITION 6: The **Multi-entry Inter-coupled Attribute Value Similarity (MeIeAVS)** between value x and y of an attribute j is:

$$\delta_j^{I_e}(x,y) = \sum_{k=1, k\neq j}^n \alpha_k \delta_{j|k}(x,y), \qquad (10)$$

In the experiment, we set the weight α_k for each attribute to a same value.

In the toy example, as attribute a_2 has multiple entries, we get $\delta_{2|3}^I(C_1, C_2) = 0.64$. Meanwhile, since a_3 is a single-entry attribute, the calculation of IRS $\delta_{3|2}^I(B_1, B_2)$ follows the original single-entry formulation, the result is 0.65.

DEFINITION 7: The **Multi-entry Coupled Attribute VAlue Similarity (MeCAVaS)** $\delta^{Aa}(x,y)$ between attribute values x and y of an attribute a_j is the product of intra-coupled and inter-coupled similarity.

$$\delta_j^{Aa}(x,y) = \delta_j^{I_a}(x,y) \cdot \delta_j^{I_e}(x,y), \qquad (11)$$

E. Coupled Attribute Vector Similarity

DEFINITION 8: For a multi-entry attribute a_k, we define the **Multi-entry Coupled Attribute VEctor Similarity (MeCAVeS)** between attribute vectors $f_k(u_i)$ and $f_k(u_j)$ of two objects u_i and u_j as $\delta_k^{Ae}(f_k(u_i), f_k(u_j))$.

Before designing the multi-entry vector similarity, we introduce some desirable characteristics at first.

1) Sparsity inequality: Rarely occurred attribute values usually convey larger discriminative power. Suppose we have got two images both tagged by "peach". Apparently, they are more similar than two "fruit" images.

$$\delta^{Ae}(\{w_i\}, \{w_i\}) < \delta^{Ae}(\{w_j\}, \{w_j\}), \text{ if } \mathcal{F}_{\{w_i\}}(U) > \mathcal{F}_{\{w_j\}}(U). \quad (12)$$

2) Aggregation inequality: Objects with more same values are more similar to each other.

$$\delta^{Ae}(\{w_i\}, \{w_i\}) < \delta^{Ae}(\{w_i, w_j\}, \{w_i, w_j\}) \quad (13)$$

3) Full profile tendency: It is an optional characteristic that we never penalize additional attribute values. The aim is to encourage images with more tags. Tag-rich images tend to receive higher similarity to other images. Therefore, in the tagging application, the system can get more candidate tags.

$$\delta^{Ae}(\{w_i\}, \{w_i\}) \leq \delta^{Ae}(\{w_i, w_{j_1}, ..., w_{j_p}\}, \{w_i, w_{k_1}, ..., w_{k_q}\}) \quad (14)$$

where $\{J\} \bigcap \{K\} = \emptyset$.

To achieve the sparsity feature, we calculate the abstraction metric of each attribute value. Concepts with higher abstraction metric get lower weight when calculating the similarity metric.

DEFINITION 9: Given the information sub-table S_k, we define in-degree, out-degree and abstraction metric of an attribute value x as:

$$D_k^{in}(x) = \frac{\sum_j (P((j, x)|j))}{|V_k| - 1}, j \in V_k \setminus \{x\} \quad (15)$$

$$D_k^{out}(x) = \frac{\sum_j (P((j, x)|x))}{|V_k| - 1}, j \in V_k \setminus \{x\} \quad (16)$$

$$Abst_k(x) = D_k^{out}(x)/D_k^{in}(x) \quad (17)$$

We propose a championship list selection strategy to turn attribute value similarity into vector similarity. Given two attribute vectors, the algorithm iteratively find the closest attribute value from each vector, shrink the vectors, and update the vector similarity according to value similarity and value abstraction. The details are shown is Algorithm 1.

Algorithm 1 Championship list selection algorithm

Input: Attribute vectors for two objects $f_k(u_i)$ and $f_k(u_j) \subseteq V_k$. MeCAVaS $\delta^{Aa}(x, y)$ for each $x, y \in T$
Output: MeCAVes $\delta^{Ae}(f_k(u_i), f_k(u_j))$.

   initialize $\delta^{Ae} = 0$
   repeat
      retrieve an attribute value from each object $w_i^{(t)}$ and $w_j^{(t)}$ who get the
      minimum $\delta^{Aa}(w_i^{(t)}, w_j^{(t)})$
      let $s = \max\{Abst_k(w_i^{(t)}), Abst_k(w_i^{(t)})\}$
      update $\delta^{Ae} += \delta^{Aa}(w_i^{(t)}, w_j^{(t)}))/s$
      remove $w_i^{(t)}$ from $f_k(u_i)$ and $w_j^{(t)}$ from $f_k(u_j)$
   until $f_k(u_i)$ or $f_k(u_j)$ is empty

In the toy example, $\delta_1^{Ae}(f_{a_2}(u_2), f_{a_2}(u_6)) = \delta_2^{Aa}(B_1, B_1)/Abst_2(B_1) + \delta_2^{Aa}(B_2, B_3)/Abst_2(B_3) = 0.625/1.1 + 0.031/1.4 = 0.59$.

F. Coupled Object Similarity

Given the information table S, the **Coupled Object Similarity (COS)** between objects u_{i_1} and u_{i_2} is:

$$COS(u_i, u_j) = \sum_{k=1}^{n} \delta_k^A(f_k(u_i), f_k(u_j))/n \quad (18)$$

where

$$\delta_k^A = \begin{cases} \delta_k^{Aa} & \text{j is a single-entry attribute} \\ \delta_k^{Ae} & \text{j is a multi-entry attribute} \end{cases}$$

where n is the number of attributes.

Finally, for COS, all CAVSs of single-entry and multi-entry attributes are added together. Therefore, we have accomplished the modified coupled similarity algorithm starting from the similarity between attribute values to the similarity between objects.

In the toy example, $COS(u_1, u_2) = (0.125 + 0.285 + 0.36)/3 = 0.256$.

V. DATASET

We present experiments on a Flickr Group dataset. Groups in Flickr are self-organized communities with common interests [34]. People participate in groups by sharing photos or discussing topics for specific social demand. Analysis on Flickr groups could offer insights into robust content presentation and social behavior trends.

A. Overview

Using the Flickr API, we collect data from 15 Flickr groups covering different interests including nature, travel, art, news, and animation. We use seven of them and extract the latest 3000 images for groups with a large number of images. Tags that occur in less than five images are discarded. Table IV provides an overview of the dataset used in this study. We number them Group A to Group G in the rest of this paper.

TABLE IV. FLICKR GROUP DATASET

ID	Group Name	# Img	# Tag	# Usr
10477049@N00	The Southwest United States	3000	973	122
1063441@N20	:<VOCALOIDS>:	889	312	106
1255404@N21	Shanghai Open Gallery	3000	665	187
1314582@N21	Fascinating Nature: Level 1	3000	878	709
655343@N23	Creative Art Photography	3000	398	328
88657993@N00	Design	3000	985	554
94326334@N00	News-Photojournalism	3000	1250	359

B. Preliminary

In a social environment, images are associated with various types of attributes. In this paper, we mainly focus on visual features and three contextual features: user, tag, and time. There are several other features available on Flickr, such as location, comments and user friendship, which may also be incorporated into the proposed algorithm.
Visual Features.
(1) *Color.* We use two color based features: color histogram and color moments.
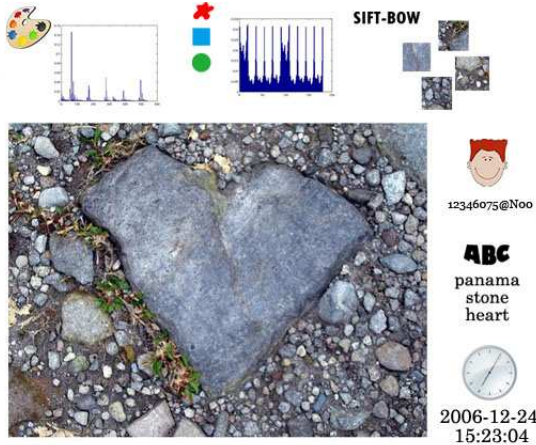
Fig. 3. Demonstration for visual contents and metadata contexts for an image.

(2) *Texture.* Local Binary Patterns (LBP) is used to extract texture features [35].

(3) *Interest Points.* We extract SIFT features to provide scale and orientation invariance [36]. SIFT is known as the most successful feature descriptor in image classification. SIFT descriptors are then quantified into visual words by bag-of-words model [37]. Afterwards, bag-of-words model is performed to each of the features to obtain categorical attribute values. We extract 50 words for each visual feature.

Contextual Metadata.

(1) *User.* In Flickr, images are uploaded by users freely. Thus, the ownership conveys important contextual information. An image is linked to its owner. The user-image relationship is one-to-one.

(2) *Tag.* Tags are descriptive texts assigned by image owners to describe the semantics of images. Tags indicate both visual information ("sky", "blue") and semantic information ("happy", "canon", "awesome"). Tags may be the most important attribute to connect images to their semantics. An image can have zero to thousands of tags, the relationship between tags and images is many-to-one.

(3) *Time.* An image in Flickr is associated with a timestamp indicating its uploading time.

VI. EXPERIMENTS

In this section, we present three typical applications of the multi-entry coupled object similarity (MeCOS) metrics[1]. Using the dataset of Flickr group (described in Section V), we first present a brief analysis of the tag network generated according to the multi-entry coupled attribute value similarity (MeCAVS). Section VI-B presents case studies on extracting relational clusters driven by local image similarity. A quantitative evaluation of the algorithm through a tag prediction task is provided in Section VI-C. The performance of the algorithm is compared with baseline methods without introducing coupling

---

[1]For detailed results, please visit http://ir.sjtu.edu.cn/~zxu/cp.html.

relationship. We also present a comparison on different applications with another algorithm that discovers multirelational structure in social images.

A. Tag Similarity

By taking tags as an attribute of images, we use the intermediate result (MeCAVS in Section IV-E) to measure the similarity between tags. Fig. 4 shows the tag network generated from the top 100 tags in the group "The Southwest United States".

Compared to the result obtained by tag co-occurrence, the proposed algorithm removes fake links generated by the most frequently occurred tags which may be treated as stopwords in document processing. As a result, the network is much clearer and meaningful. The network contains links that appears naturally (e.g. "nm" v.s. "newmexico") and ones generated according to specific dataset (e.g. "coloradowildlife" v.s. "deer"). In the coupled view, except for the tag co-occurrence, other factors including users and visual contents also play an important role in measuring the tag similarity.

B. Image Clustering

One of the most important applications of the proposed algorithm is to extract relational clusters from a group image stream. In this application, we present an effective algorithm to organize contents in a Flickr group. We cluster the images into a given number of patches. In each patch, the most representative images, tags, users and other useful information are presented, providing a user-friendly demonstration for the group.

1) Experimental Settings: Specifically, each image in the group is considered as an object, whose attributes include tag, user and visual features. Object similarity is calculated by MeCOS algorithm. We choose k-modes [39] as the clustering algorithm in this study, which can deal with the problem of categorical data clustering.

To evaluate the clustering result, we seek to answer the following questions.

- What are the most representative images, tags and users in each cluster?
- To what extent do the clusters separate from each other?
- Which attribute plays the most important role in the clustering progress?
- Does the clustering result really help organizing multi-relational data and exploring images?
- Does the algorithm obtain similar results on different groups?

We enclose clustering results of groups driven by different interests. In each group, we extract 5 relational clusters and show the most representative tags and images. In fact, images in the mode sequence extracted by k-modes algorithm are in the centroid of the cluster images. Therefore, they can be directly selected as the representative images. Here the algorithm extracts 5 images in the mode sequence for each cluster.

Fig. 4. Tag network generated by MeCAVS for group "The Southwest United States", drawn using NetDraw software [38]. The figure shows all links between the top 100 frequently occurred tags with MeCAVS value larger than 0.1. Different colors stand for different components extracted from the tag map. Relationship shown in the figure is much clearer than that from tag co-occurrence.

2) Clustering Results: Fig. 5 shows the clustering result of the group "Creative Art Photography", which collects images showing creative minds. The collections contain paintings, photographs and post-processed images. From the most popular tags we know that the tags mostly concentrate on content("flower", "water"), quality("supershot", "diamondclassphotographer"), camera information("nikon") and artistic tricks("macro", "photoshop"). Cluster A1 mainly collects photos describing portrait and creative activities. A significant feature is that there are many black-and-white photographs in the cluster. The other big cluster, Cluster A3 consists of highquality macro shots of flowers and nature scenes. Cluster A4 shows colorful paintings. Images performing High-Dynamic Range(HDR) to depict clouds, sky, sunset and beach are collected in Cluster A5. Cluster A2 is a little bit different that it is characterized by tags indicating high quality (e.g."supershot", "bravo"). In general, the clustering process is mainly guided by visual features. It is clear that by utilizing the multi-relational information, the algorithm successfully extracts interpretable structure of the data. By viewing the group again in a better organized view, we can summarize the group as a community that shares high-quality shots of various semantic concepts (e.g. portraits, flowers, trees and sunset) utilizing creative artistic tricks (e.g. black-and-white, HDR and macro-shot).

In Fig. 6(a), analysis on a news event oriented group ("News-Photojournalism") shows a different characteristic. As most of the contents are describing local events based on abstract concepts, visual features play a weaker role in the clustering process. We discover that the representative tags of five clusters all tell a story, indicating five news events respectively (politics demonstration, Iraq war, the Bush government, poverty problem in the Philippines and a fire in Pottstown). We apply the algorithm to other groups and achieves analogous results. The results show that the proposed algorithm organizes the group content in a more interpretable way. Meanwhile, based on the relational structure, local events can be extracted effectively.

3) Comparison with Global-based Algorithm: As mentioned in Section I, Lin et al. [4] proposed a nonnegative matrix factorization algorithm (shorted to NMF in below) to characterize relational structures of group photo streams. The basic idea is to minimize the difference between observation data matrix and predicting matrix (a product of three component matrices). We present a comparison with their clustering results. Since the evaluation is a largely subjective issue, the discussion mainly focus on the merits and shortcomings of global and local based clustering methods.

The global-based matrix factorization algorithm achieves nearly optimized clustering results on singular attributes. With a pairwise comparison, we find that two methods tends to extract similar news events. However, as shown in Fig. 6(b), the method suffers a problem that the representative images or tags are not "representative" enough. We recall that in [4], the authors select the most representative images based on how likely an image i belongs to a cluster k. It is argued that even an image i has a 100% probability for cluster k, it is not convincing that the image is the most suitable choice for representing the cluster. Experimental results show that the representative images selected by the proposed algorithm as the modes in k-modes algorithm are more interpretable than those selected by NMF method.

In addition, the proposed algorithm has several features that outperforms the global-based algorithm.

Image relationship map. With the similarities between all image pairs, the system can extract not only the cluster membership for each image, but also the structure within and between clusters. For example, we can pick out the centroids, outliers within a cluster, along with the bridge nodes between two clusters. It gives a better understanding of the relational structure of the data.

New item problem. When a new image is added into the group, the global-based algorithms need to re-calculate the whole matrix to reach the global minimum point. In contrast, for the proposed algorithm, the only additional calculation is
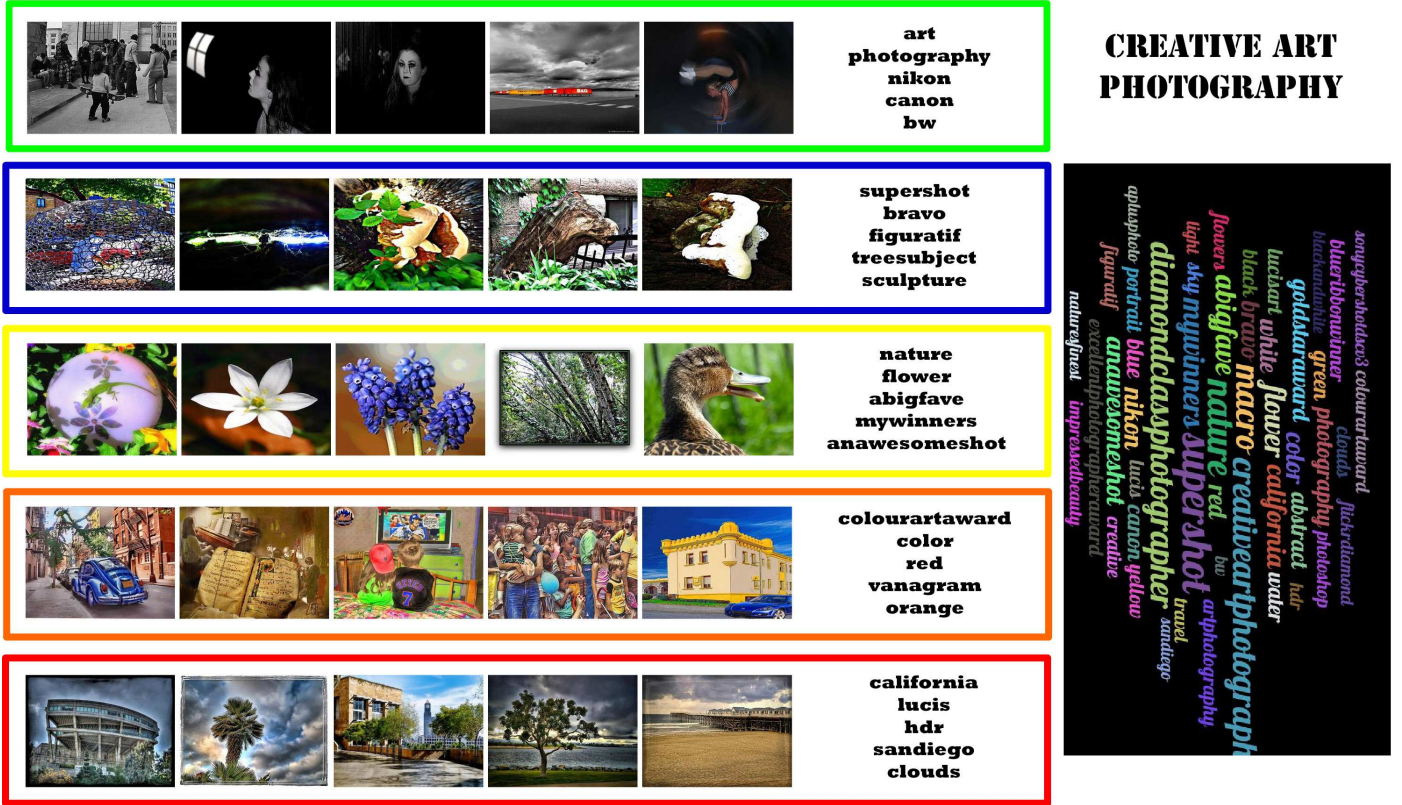
Fig. 5. Clustering result of the Group "Creative Art Photography". The tagCloud generated by the top 100 tags [40] is shown in the right part. In the right bottom shows the image-tag and image-user matrices where the rows and columns are reordered based on the cluster indices of the corresponding images, tags or users. The most representative five images and tags of each cluster are presented in the left. Images are resized into the same size. The tags and images of clusters indicate strong semantic information.

to measure the similarity between the new image and every image in the dataset. The updating process is much easier and faster.

C. Tag Recommendation

Tag prediction task provides a quantitative evaluation for the proposed similarity metrics. We perform nearest neighbour algorithm with MeCOS as the similarity metric. The underlying assumption is that visually similar images tend to have similar tags, especially ones uploaded by users sharing a common taste. It is worthwhile to note that given the pairwise image similarity, more complicated diffusion methods can potentially be applied to obtain better prediction performance.

For each group, we choose 80% images for training and 20% for testing. Flickr has an important strategy named batch tagging. A user can tag dozens of images using a batch mode when uploading them. If a batch of images happen to appear in testing set and training set at the same time, predicting by the tags of these training images will lead to a fake high performance. Therefore, instead of random choosing, we split the training and testing set according to the uploading timestamp.

We compare our result with NMF and two baseline methods based on tag frequency and visual similarity relatively. We use
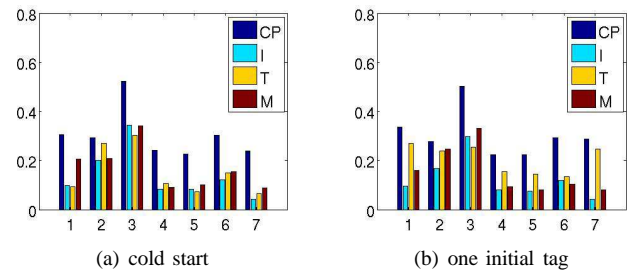

Fig. 7. P@10 for seven Flickr groups. CP stands for coupled similarity metric, I stands for image feature similarity, T stands for top occurrence tags and M stands for NMF. The left one shows the result of cold start, while the right one is the result of given one initial tag as input.

P@10 (precision of the top 10 predicted tags) as the evaluation metric. Experimental results show that other criterions such as R@10 (recall of the top 10 predicted tags), NDCG (Normalized Discount Cumulative Gain) get similar results.

The empirical results show that the proposed algorithm outperforms baseline algorithms significantly. For groups whose tags are averagely scattered, NMF outperforms tag-based method. However, for the groups which have several
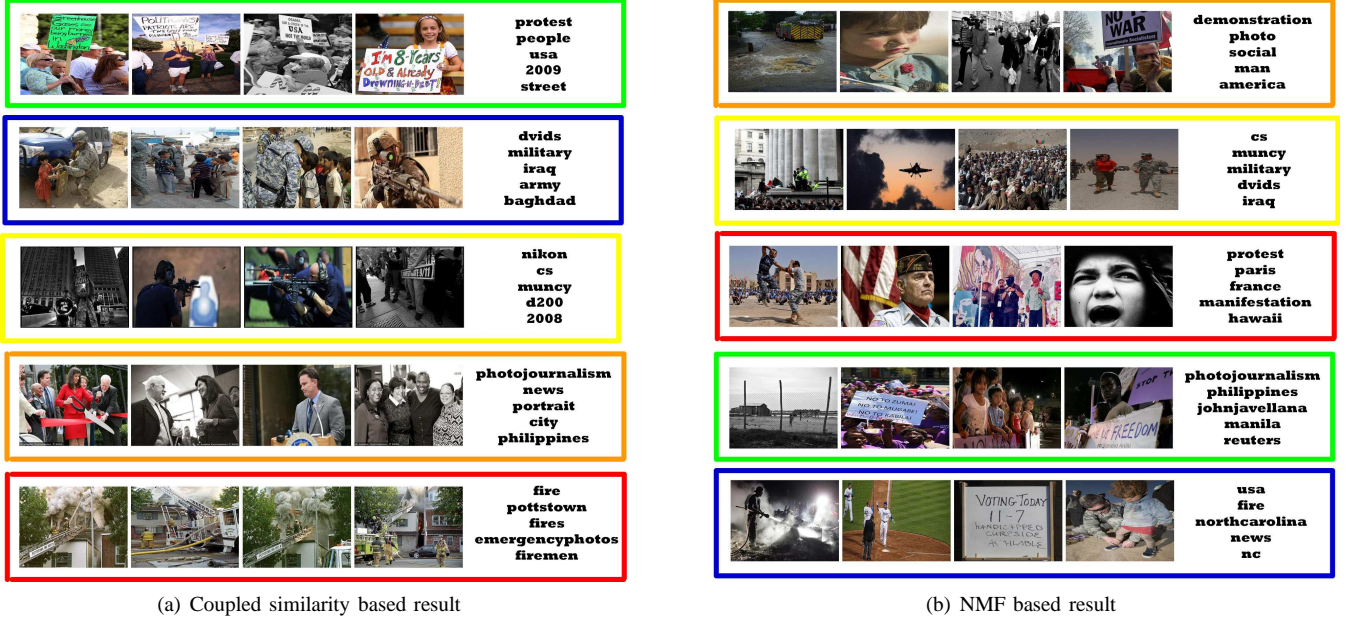
(a) Coupled similarity based result

(b) NMF based result

Fig. 6. Clustering result of the Group "News-Photojournalism". (a) The result generated by MeCOS algorithm. The result shows an event-centric structure. Visual features play a weaker role in the clustering process. (b) The result generated by NMF. Clusters are re-ordered to enable pairwise comparison to coupled algorithm. Although NMF method extracts reasonable events, the representative images are not clear enough according to the clusters' representative tags.

group related tags, baseline methods achieves competitive or even better performance than NMF. It is argued that NMF, as a global-based method, uses the same weight for all images in the cluster when predicting tags. Hence, images that are less similar to the predicted image could produce negative impact on the predicting performance. On the contrary, MeCOS based algorithm only gives high weight for the most similar images.

When evaluating the performance of the proposed algorithm in predicting tags for images, all tags of testing images are removed. Therefore, tags are recommended to testing images without knowing their initial tags. In the view of recommender system, this phenomenon is formulated as *cold start problem*. When we introduce an experiment on performing collaborative filtering algorithm on the dataset, result shows that for every image, the most frequently occurred tags are recommended by the algorithm without personalized variance.

We present a performance evaluation to measure the performance of different startup parameters. For every testing image, the first N tags are given as initial input, leaving the remaining tags as the predicting goal. In the first step, no initial tags are given, similar images are ones which are visually similar and uploaded by similar users. When N increases, with more tags given as input, algorithms tend to find images that are more similar in semantic level. For the tag frequency method, we added a Naive Bayes algorithm given the tag co-occurrence matrix. Experiments (Fig. 8) show that although NMF and tag-based method both increase their performance when given more input tags, MeCOS algorithm still outperforms other methods constantly.

A typical demonstration of different predicting results when importing new tags is shown in Fig. 9. The recommending images become more similar in semantic level than visual level with the raise of input tags.

VII. DISCUSSIONS

We present three applications of MeCOS in the paper. By introducing coupling relationship of multiple attributes, MeCOS shows promising performance on non-IID social images. The result is based on the following characteristics of MeCOS.

Multiple features. Compared to baseline methods such as image similarity and tag similarity, MeCOS employs multiple attributes available for social images to achieve better performance.

Coupling relationship. Instead of regarding each attribute independently, MeCOS introduces coupling relationship between multiple attributes. The inter-coupled relationship provides valuable information in addition to traditional attribute value similarity.

Interpretable results. There are some algorithms that employ inter-relationship between attributes, such as biclustering [24], hypergraph [3] and topic model [22]. However, most of them are not transparent in the training process, which makes the result hard to explain. On the contrary, MeCOS provides a clearer explanation of why two objects are similar by tracing the similarity between multiple attribute values. This characteristic is extremely important for recommender systems. When introducing possible friends for a user, a higher success rate will be achieved if a plausible reason is provided.

Versatility. By placing all the attributes on a same status, MeCOS provides a unified view to analyze social images. Having attained all the inter-coupled and intra-coupled similarity of each attribute, the algorithm can be extended to analyze

(a) P@10 for group A

(b) P@10 for group B

(c) P@10 for group D

(d) P@10 for group E
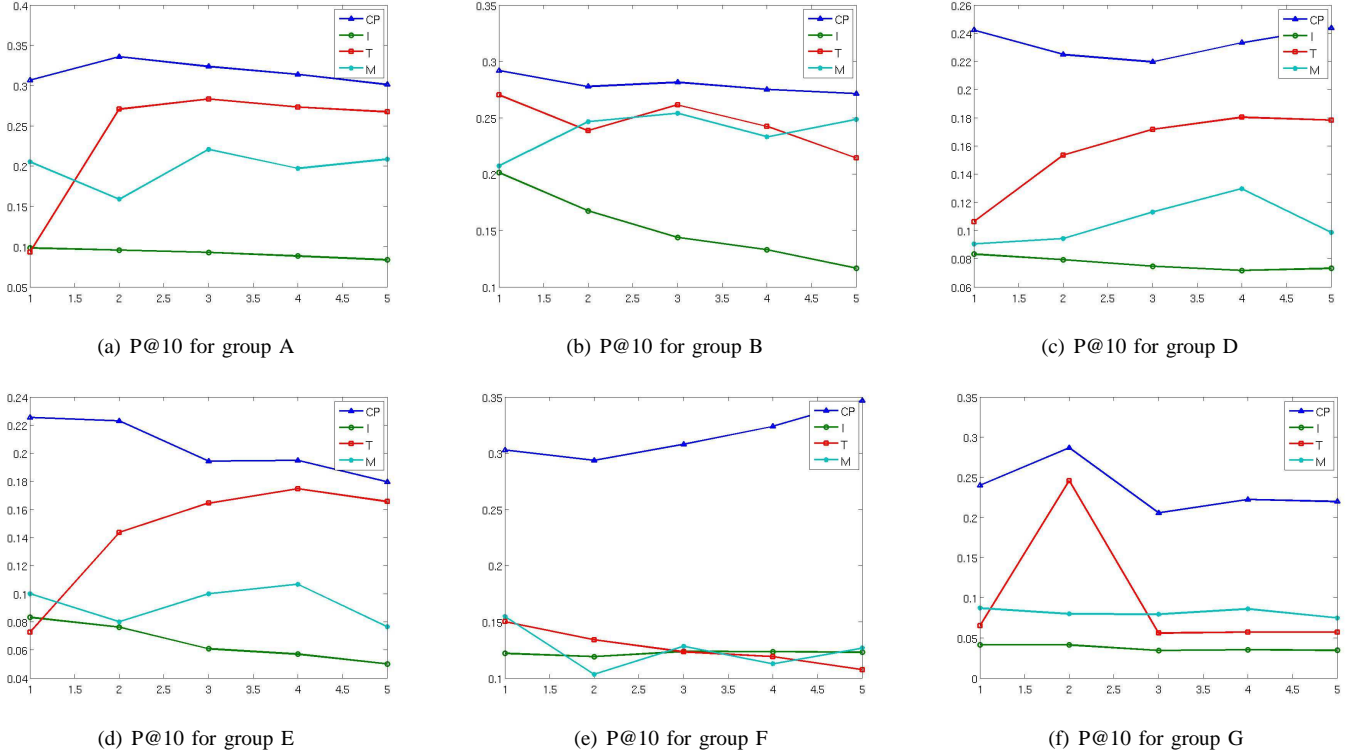
(e) P@10 for group F

(f) P@10 for group G

Fig. 8. Cold start result for different groups on P@10. The x-axis stands for number of input tags. Coupled based algorithms (CP) outperforms image similarity (I), popular tags (T) and NMF (M) methods. Note that with the increase of input tags, the total number of tags available in the testing set declines. Therefore, the upper bound of P@10 declines.

all problems related to the concept "similarity". Although the performance may not be competitive with state-of-the-art algorithms which are designed specifically for one problem, MeCOS does provide a new perspective of social image analysis.

Despite the merits above, MeCOS still has much room for improvement. The main problem of MeCOS is that when computing coupled object similarity, the weight of each attribute is fixed. Therefore, since Flickr Group dataset is highly determined on the role of users, the clustering result of MeCOS is very sensitive to users. In that case, the weight of user attribute should be lower. A set of self-learned parameters is preferred for different datasets.

For future work, we plan to extend the application of the algorithm to efficient image discovery and aesthetic image analysis. We also plan to build an online application demo to provide a better demonstration for the structure of contents in multimedia social sharing websites.

VIII. CONCLUSION

Fully capturing relationships between social images in social applications including social media is very important but challenging. The challenge lies in the deep understanding of comprehensive coupling relationships between images, social image attributes and values. This makes social images essen-

tially non-IID. This motivates us to explore theoretical foundation for analyzing non-IID social images in this paper. Accordingly, we present a new and effective similarity metric: multi-entry coupled similarity metrics and social image analyzing algorithm to capture the comprehensive couplings embedded in social images and especially the many-to-one relationship between tags and images. By measuring the similarity between images together with the similarity within various kinds of contexts at the same time, we provide a unified view for tackling social image applications including network analysis, image clustering and tag prediction. Experiments show that the algorithm discovers interpretable structures of the multimedia contents especially in terms of representative images and tags, which provides a more user-friendly demonstration of social images. Extensive experiments show that the proposed similarity metric and algorithm outperform baseline methods and global based matrix factorization method in tag prediction.

REFERENCES

[1] X.-J. Wang, L. Zhang, F. Jing, and W.-Y. Ma, "Annosearch: Image auto-annotation by search," in *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, vol. 2. IEEE, 2006, pp. 1483–1490.

[2] L. Wu, X.-S. Hua, N. Yu, W.-Y. Ma, and S. Li, "Flickr distance," in *Proceedings of the 16th ACM international conference on Multimedia*. ACM, 2008, pp. 31–40.

Fig. 9. Top 10 similar images retrieved by different algorithms and parameters. From top to bottom: coupled similarity metric cold start level 0-4, image similarity, tag similarity. All images are resized to the same size. In the left shows the added tag in each round. The left bottom image is the testing image.

[3] F. Wu, Y.-H. Han, and Y.-T. Zhuang, "Multiple hypergraph clustering of web images by miningword2image correlations," *Journal of Computer Science and Technology*, vol. 25, no. 4, pp. 750–760, 2010.

[4] Y.-R. Lin, H. Sundaram, M. De Choudhury, and A. Kelliher, "Discovering multirelational structure in social media streams," *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP)*, vol. 8, no. 1, p. 4, 2012.

[5] L. Cao, "Non-iidness learning in behavioral and social data," *The Computer Journal*, 2013.

[6] C. Wang, L. Cao, M. Wang, J. Li, W. Wei, and Y. Ou, "Coupled nominal similarity in unsupervised learning," in *Proceedings of the 20th ACM international conference on Information and knowledge management*. ACM, 2011, pp. 973–978.

[7] Y. Yu, C. Wang, Y. Gao, L. Cao, and X. Chen, "A coupled clustering approach for items recommendation," in *Advances in Knowledge Discovery and Data Mining*. Springer, 2013, pp. 365–376.

[8] H. Becker, M. Naaman, and L. Gravano, "Learning similarity metrics for event identification in social media," in *Proceedings of the third ACM international conference on Web search and data mining*. ACM, 2010, pp. 291–300.

[9] S. Papadopoulos, C. Zigkolis, S. Kapiris, Y. Kompatsiaris, and A. Vakali, "Clusttour: city exploration by use of hybrid photo clustering," in *Proceedings of the international conference on Multimedia*. ACM, 2010, pp. 1617–1620.

[10] H. Zhang, M. Korayem, D. J. Crandall, and G. LeBuhn, "Mining photo-sharing websites to study ecological phenomena," in *Proceedings of the 21st international conference on World Wide Web*. ACM, 2012, pp. 749–758.

[11] K. Lerman, A. Plangprasopchok, and C. Wong, "Personalizing image search results on flickr," in *American Association for Artificial Intelligence*, 2007.

[12] N. Garg and I. Weber, "Personalized, interactive tag recommendation for flickr," in *Proceedings of the 2008 ACM conference on Recommender systems*. ACM, 2008, pp. 67–74.

[13] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 22, no. 12, pp. 1349–1380, 2000.

[14] B. Sigurbjörnsson and R. Van Zwol, "Flickr tag recommendation based on collective knowledge," in *Proceedings of the 17th international conference on World Wide Web*. ACM, 2008, pp. 327–336.

[15] C. Wang, L. Zhang, and H.-J. Zhang, "Learning to reduce the semantic gap in web image retrieval and annotation," in *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*. ACM, 2008, pp. 355–362.

[16] O. Nov, M. Naaman, and C. Ye, "What drives content tagging: the case of photos on flickr," in *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2008, pp. 1097–1100.

[17] L. Cao, J. Yu, J. Luo, and T. S. Huang, "Enhancing semantic and geographic annotation of web images via logistic canonical correlation regression," in *Proceedings of the 17th ACM international conference on Multimedia*. ACM, 2009, pp. 125–134.

[18] I. Simon, N. Snavely, and S. M. Seitz, "Scene summarization for online image collections," in *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*. IEEE, 2007, pp. 1–8.

[19] P. Serdyukov, V. Murdock, and R. Van Zwol, "Placing flickr photos on a map," in *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*. ACM, 2009, pp. 484–491.

[20] G. Palla, A.-L. Barabási, and T. Vicsek, "Quantifying social group evolution," *Nature*, vol. 446, no. 7136, pp. 664–667, 2007.

[21] R. Kumar, J. Novak, and A. Tomkins, "Structure and evolution of online social networks," in *Link Mining: Models, Algorithms, and Applications*. Springer, 2010, pp. 337–357.

[22] D. M. Blei and J. D. Lafferty, "Dynamic topic models," in *Proceedings of the 23rd international conference on Machine learning*. ACM, 2006, pp. 113–120.

[23] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *the Journal of machine Learning research*, vol. 3, pp. 993–1022, 2003.

[24] X. Rui, M. Li, Z. Li, W.-Y. Ma, and N. Yu, "Bipartite graph reinforcement model for web image annotation," in *Proceedings of the 15th international conference on Multimedia*. ACM, 2007, pp. 585–594.

[25] L. Tang, H. Liu, J. Zhang, and Z. Nazeri, "Community evolution in dynamic multi-mode networks," in *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2008, pp. 677–685.

[26] A. Banerjee, S. Basu, and S. Merugu, "Multi-way clustering on relation graphs," *SDM. SIAM*, 2007.

[27] Y. Sun, J. Han, X. Yan, P. S. Yu, and T. Wu, "Pathsim: Meta path-based top-k similarity search in heterogeneous information networks," *VLDB11*, 2011.

[28] L. Cao, Y. Ou, P. S. YU, and G. Wei, "Detecting abnormal coupled

sequences and sequence changes in group-based manipulative trading behaviors," in *KDD2010*, 2010, pp. 85–94.

[29] L. Cao, Y. Ou, and P. S. Yu., "Coupled behavior analysis with applications," *IEEE Trans. on Knowledge and Data Engineering*, vol. 24(8), pp. 1378–1392, 2012.

[30] Y. Song, L. Cao, X. Wu, G. Wei, W. Ye, and W. Ding, "Coupled behavior analysis for capturing coupling relationships in group-based market manipulations," in *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2012, pp. 976–984.

[31] F. Li, G. Xu, L. Cao, and Z. Niu., "Coupled group-based matrix factorization for recommender system," in *WISE*, 2013.

[32] C. Wang, Z. She, and L. Cao, "Coupled clustering ensemble: Incorporating coupling relationships both between base clusterings and objects," in *Data Engineering (ICDE), 2013 IEEE 29th International Conference on*. IEEE, 2013, pp. 374–385.

[33] G. Gan, C. Ma, and J. Wu, *Data clustering: theory, algorithms, and applications*. Society for Industrial and Applied Mathematics, 2007, vol. 20.

[34] R. A. Negoescu and D. Gatica-Perez, "Analyzing flickr groups," in *Proceedings of the 2008 international conference on Content-based image and video retrieval*. ACM, 2008, pp. 417–426.

[35] L. Wang and D.-C. He, "Texture classification using texture spectrum," *Pattern Recognition*, vol. 23, no. 8, pp. 905–910, 1990.

[36] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, pp. 91–110, 2004, 10.1023/B:VISI.0000029664.99615.94. [Online]. Available: http://dx.doi.org/10.1023/B:VISI.0000029664.99615.94

[37] J. Sivic and A. Zisserman, "Efficient visual search of videos cast as text retrieval," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 31, no. 4, pp. 591–606, 2009.

[38] S. Borgatti, "Netdraw software for network visualization," *Analytic Technologies*, 2002.

[39] Z. Huang, "Extensions to the k-means algorithm for clustering large data sets with categorical values," *Data Mining and Knowledge Discovery*, vol. 2, no. 3, pp. 283–304, 1998.

[40] M. J. Halvey and M. T. Keane, "An assessment of tag presentation techniques," in *Proceedings of the 16th international conference on World Wide Web*. ACM, 2007, pp. 1313–1314.