Visualizing Brand Associations from Web Community Photos

Gunhee Kim
Disney Research
Pittsburgh, 15213 PA
gunhee@cs.cmu.edu

Eric P. Xing
Carnegie Mellon University
Pittsburgh, 15213 PA
epxing@cs.cmu.edu

ABSTRACT

Brand Associations, one of central concepts in marketing, describe customers' top-of-mind attitudes or feelings toward a brand. (*e.g.* what comes to mind when you think of *Nike*?) Thus, this consumer-driven brand equity often attains the grounds for purchasing products or services of the brand. Traditionally, brand associations are measured by analyzing the text data from consumers' responses to the survey or their online conversation logs. In this paper, we propose to leverage large-scale online photo collections contributed by the general public, given that photos are gaining popularity as an important information modality on the Web. As a first technical step toward photo-based brand association study, we aim to jointly achieve the following two visualization tasks in a mutually-rewarding way: (i) detecting and visualizing core visual concepts associated with brands, and (ii) localizing the regions of brand in images. With experiments on about five millions of images of 48 brands crawled from five popular online photo sharing sites, we demonstrate that our approach can discover complementary views on the brand associations that are hardly mined from text data. We also quantitatively show that our approach outperforms other candidate methods on both visualization tasks.

Categories and Subject Descriptors

I.4.9 [**Image processing and computer vision**]: Applications; J.4 [**Computer Applications**]: Social and behavioral sciences—*Economics*

Keywords

Brand associations, summarization and visualization of multimedia data, image segmentation

1. INTRODUCTION

Brand equity describes a set of values or assets linked to a brand [1, 13]. It is one of core concepts in marketing since it is a key source of bearing the competitive advantage of a company over its competitors, boosting efficiency and effectiveness of marketing programs, and attaining the price premium due to increased customer satisfaction and loyalty, to name a few. A central component of brand equity is *brand associations*, which are the set of associations that consumers perceive with the brand [13]. For example, the brand associations of *Nike* may include *Tiger Woods*, *shoes*, and *basketball*. Its significance lies in that it is a *customer-driven* brand equity; that is, the brand associations are directly connected to customers' *top-of-mind* attitudes or feelings toward the brand, which provoke the reasons to preferentially purchase the products or services of the brand. For instance, if a customer strongly associates *Nike* with *golf shirts*, he may tend to first consider *Nike* products over other competitors' ones when he needs one.

Traditionally, measuring brand associations is a challenging task because it should be built from direct consumer responses to carefully designed questionnaires [2, 5, 24, 26]. With the recent emergence of online social media, it has been developed to indirectly leverage consumer-generated data on online communities such as Weblogs, boards, and Wiki. Beneficially, resources on such social media are obtainable inexpensively and almost instantaneously from a large crowd of potential customers. One typical example of such practice is the *Brand Association Map* developed by *Nielsen Online* [2, 19], in which important concepts and themes correlated with a given brand name are automatically extracted from the data of online conversations.

In this paper, for the study of brand associations, we propose to go beyond textual media, and take advantage of large-scale online *photo* collections, which have not been explored so far. Admittedly, pictures can be inferior to mine subjective sentiments than texts (*e.g. Nike is too expensive*). However, pictures can be a complementary information modality to show customers' experiences regarding brands within a natural context. With widespread availability of digital cameras and smartphones, people can freely take pictures on any memorable moments, which include experiencing or purchasing products they like. In addition, many online tools enable people to easily share, comment, or bookmark the images of products that they wish to buy.

A complete solution to photo-based brand associations can be too challenging to be achieved in a single paper since it requires not only technical algorithms but also sophisticated user interfaces for marketers and general users based upon thorough user studies. Thus, as a first technical step, in this paper, we address the problem of jointly achieving the following two levels of visualization tasks regarding brand

WSDM '14 New York City, U.S.A

(a) Brand association map (b) Brand localization

Figure 1: Motivation for two visualization tasks toward brand association study from Web community photos with two competing brands of *Nike* and *Adidas*. (a) Task1: we perform exemplar detection and clustering to reconstruct brand association maps (BAM). A more strongly associated cluster with the brand appears closer to the center of the map. A higher correlated pair of clusters has a smaller angular distance. We show top 20 exemplars (*i.e.* cluster centers) in the map. On the right, for some selected exemplars, we show the average images of 40 nearest neighbors in their corresponding clusters. (b) Task2: we segment the most likely regions of brand in the images.

associations. (See the examples in Fig.1).

(1) *Visualizing core pictorial concepts associated with brands*: It has been a key problem in brand association research to concisely visualize important concepts associated with brands in a form of networks or maps [2, 6, 24, 26]. Therefore, our first task is, as shown in Fig.1.(a), to visualize core visual concepts of brands by summarizing online photos that are tagged and organized by general users. This goal involves three sub-problems: identifying a small number of image clusters and exemplars (*i.e.* cluster centers), discovering the similarity relations between clusters, and projecting them into a low-dimensional space.

(2) *Localizing the regions of brand in images*: Our second task is the *sub-image level* visualization of brand associations, while the first task addresses the *image-level* one. We aim to localize the regions that are most associated with the brand in each image in an unsupervised way (*i.e.* without any pre-defined models), as shown in Fig.1.(b). In our algorithm, we perform pixel-level image segmentation to delineate the regions of brand. Even though bounding boxes may be better as the final output to the general users, they can be trivially derived from segmentation results, by defining the minimum rectangle that encloses the segment while ignoring tiny unconnected dots.

We choose the above two tasks as the most fundamental building blocks for the study of photo-based brand associations for following reasons. The first task can provide a structural summary of large-scale and ever-growing online image data of brands, which otherwise are too overwhelming for human to grasp any underlying big picture. The second task can not only suppress background clutters, but also help reveal typical interactions between users and products in natural social scenes, which can lead a wide variety of potential benefits, ranging from content-based image retrieval to online multimedia advertisement.

Importantly, jointly solving these two tasks are *mutually rewarding*. The exemplar detection/clustering can group similar images, which can promote the brand localization since we can leverage the recurring foreground signals. In the reverse direction, localizing brand regions can enhance the similarity measurement between images, which subsequently contributes to better exemplar detection/clustering.

For evaluation, we collect about five millions of images of 48 brands of four categories (*i.e.* *sports*, *luxury*, *beer*, and *fastfood*) from five popular photo sharing sites, including Flickr, Photobucket, deviantART, Twitpic, and Pinterest. In our experiments, we show the picture-driven brand association maps for some selected brands. We also demonstrate that our approach outperforms other candidate methods on both exemplar detection/clustering and brand localization. Finally, we compare the results of our picture-based brand associations with actual sales data of brands.

1.1 Relations to Previous work

We here introduce two lines of research that are remotely related to our work.

Measuring brand associations by free association: In almost all previous research for brand associations, the surveys on customers are the main approach to collect source data. Among many ways to conduct the survey, the *free association* procedure has been one of the simplest but often most powerful ways to profile brand associations [5, 6, 26]. In this technique, subjects are asked to freely answer their feelings and thoughts about a given brand name without any editing or censoring [18]. (*e.g.* What comes to mind when you think of *Nike*?) Our research is also based on this *free association* idea, because we view the Web photos tagged with a brand name by anonymous users as their candid pictorial impressions to the brand. Therefore, from a viewpoint of brand association research, the contribution of our work is to introduce a novel source of data for the analysis.

In this line of research, the brand association map of Nielsen Online [2, 19] is closely related to our work because both approaches explore online data. However, our work is unique in that we explore online image data, which convey complementary views on the brand associations that can be missed in texts. In addition, we localize the regions of brand in every photo, which is another novel feature of our work.

Analysis of product images: Recently, with the exploding interests in electronic commerce, computer vision techniques have widely applied to analyze product images for commercial applications. Some notable examples include the product image search and ranking [11], the logo and product detection in natural images [9, 12, 16, 23], and

Web sites	Characteristics
FLICKR/ PHOTOBUCKET	Two largest and most popular photo sharing sites in terms of volumes of photos.
PINTEREST	Image collections bookmarked by users
DEVIANTART	Various forms of artwork created by users.
TWITPIC	Photos shared via Twitter.

Table 1: Five Web sites for crawling photos.

clothing parsing in fashion photos [28]. However, the objective of our work differs in that we aim to discover and visualize the core concepts of brands from uncleaned online images, whereas most of past work has focused on detecting a fixed number of specified product models or logos in the images. Therefore, in our work, it is important to mine the visual topics that do not explicitly contain the products but reflect general public's thoughts, feelings, or experiences over the brands (*e.g.* sponsored yacht competition scenes in the *Rolex* images).

1.2 Summary of Contributions

The contributions of our work are summarized as follows:

(1) We study the problem of visualizing brand associations in both image and sub-image levels by leveraging large-scale online pictures. To the best of our knowledge, our work is the first attempt so far on such photo-based brand association analysis. Our work can provide another novel and complementary way to visualize general public's impressions or thoughts on the brands.

(2) We develop an algorithm to jointly achieve exemplar detection/clustering and brand localization in a mutually-rewarding way. In addition, we design an embedding algorithm to visualize the top exemplars in a circular layout.

(3) With experiments on about five million images of 48 brands, we have found that the proposed algorithms can comprehensively but succinctly visualize key concepts of large-scale brand image collections. We also quantitatively demonstrate that our approach outperforms other candidate methods on both visualization tasks.

2. PROBLEM FORMULATION

2.1 Image Data Crawling

Since we are interested in consumer-driven views on the brands, we use the online photos that are contributed and organized by general Web users. As source data, we crawl images from the five popular photo sharing sites in Table 1. The characteristics of the pictures on the five sites are different from one another as shown in Table 1. We exclude the GOOGLE IMAGE SEARCH because much of the pictures are originated from online shopping malls or news agencies.

We query the brand names via the built-in search engines of the above sites to search for the pictures tagged with brand names. We download all retrieved images without any filtering. We also crawl meta-data of the pictures (*e.g.* timestamps, titles, user names, texts), if available.

Fig.2 summarizes our dataset of 4,783,345 images for 48 brands, which can be classified into four categories: *luxury*, *sports*, *beer*, and *fastfood*. The number of images per brand varies much according to the popularity of the brand.

2.2 Overview of Our Approach

Fig.3 presents the overview of our approach. The input of our algorithm is a set of photos for a brand of interest. Let $\mathcal{I} = \{I_1, \cdots, I_N\}$ be the set of input images, where N is the number of images. As shown in Fig.3.(b), our first step is to build a K-nearest neighbor (KNN) graph $\mathcal{G} = (\mathcal{I}, \mathcal{E})$ in which each image is connected with its K most similar images in $\mathcal{I}$. We will present our image descriptors in section 3.1, similarity measures in section 3.2, and KNN graph construction in section 3.3.

The next step is to perform exemplar detection and clustering on the graph $\mathcal{G}$, which will be discussed in section 3.4. Its goal is to discover a small set of representative images called exemplars $\mathcal{A}(\subset \mathcal{I})$, and to partition $\mathcal{I}$ so that each image is associated with its closest exemplar, as shown in Fig.3.(c). Therefore, the clusters are the groups of contextually and visually similar images, and the exemplars are the most prototypical images of the clusters.

The clustering helps discover the coherent groups of images from extremely diverse Web images, which is beneficial to detect the regions of a brand in the images (see examples in Fig.3.(d)). In our setting, the brand localization is formulated as the problem of *cosegmentation* [14, 15], which has been actively studied in image segmentation research. Its goal is to simultaneously segment out recurring objects or foregrounds across the multiple images. Obviously, the images in the same cluster are likely to share the same themes of the brand (*e.g.* bags in Fig.3.(d)), which can be discovered by the cosegmentation approach. We summarize the procedure of cosegmentation in section 3.5.

In our closed-loop approach, the segmentation can enhance the exemplar detection/clustering by promoting a more accurate image similarity measure, which will be justified in section 3.2 with an intuitive example. Hence, after finishing the cosegmentation step, we can return to the KNN graph construction and repeat the whole algorithm again with the new segmentation-based image similarity metric.

The brand association map like Fig.1 can be constructed from the exemplar detection/clustering output. The algorithm will be presented in section 4.

3. APPROACH

3.1 Feature Extraction

For image description, we use one of common practices in recent computer vision research: the dense feature extraction with vector quantization. We densely extract two most popular features from each image: HSV color SIFT and histogram of oriented edge (HOG) feature on a regular grid at steps of 4 and 8 pixels, respectively. Then, we form 300 visual words for each feature type by applying K-means to randomly selected features. Finally, the nearest word is assigned to every node of the grid. We use publicly available codes[1] for the whole process of feature extraction.

3.2 Image Similarity Measure

One prerequisite to accurate clustering is an appropriate similarity measure between images, denoted by $\sigma : \mathcal{I} \times \mathcal{I} \rightarrow \mathbb{R}$. We assert that even imperfect segmentation helps enhance the measurement of image similarity, which can justify our closed-loop approach. Fig.4 shows a typical example, in which the two images are similar in that both include persons with glasses of *Guinness* beer. For an unsegmented

[1]The SIFT and HOG feature extraction codes are available at http://www.vlfeat.org, and at http://www.cs.brown.edu/∼pff/latent, respectively.

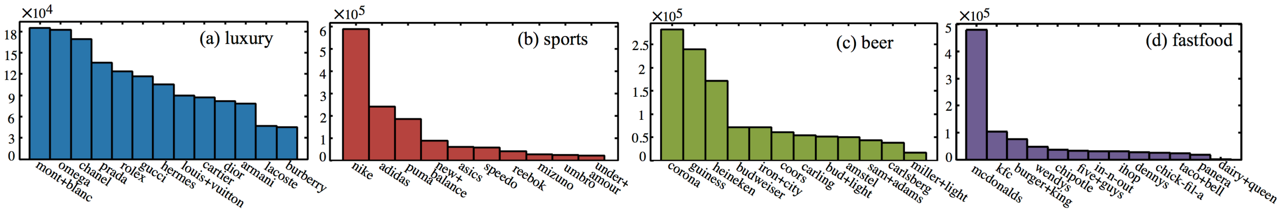

Figure 2: The dataset of 48 brands crawled from five photo sharing sites of Table 1. The brands are classified into four categories: (a) *luxury*, (b) *sports*, (c) *beer*, and (d) *fastfood*. The total number of images is 4,720,724.

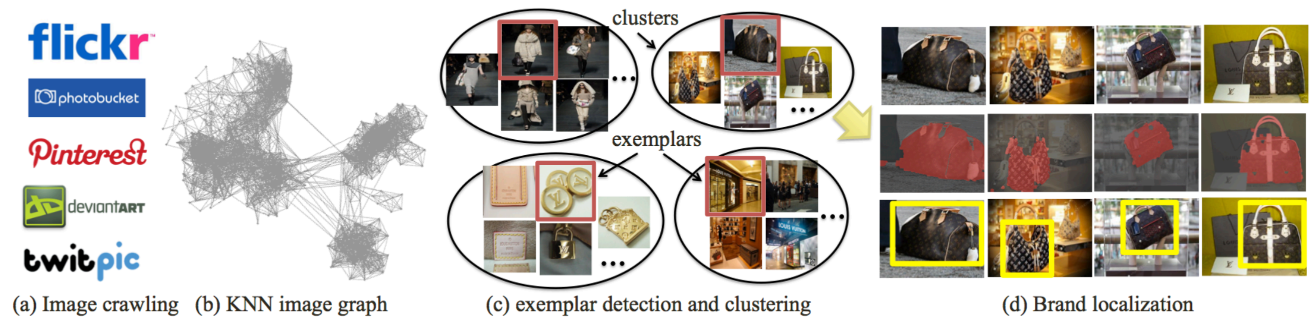

Figure 3: The overview of the proposed approach with an example of the *Louis+Vuitton*. (a) As an input, we crawl the photos of the brand from the five photo sharing sites. (b) Next, we build a K-nearest neighbor (KNN) similarity graph between images. (c) We perform the graph-based exemplar detection/clustering. (d) Finally, we cosegment the images in the same cluster in order to discover the regions of a brand in each image. As a closed-loop solution, we can return to the KNN graph construction with the new segmentation-based image similarity metric.

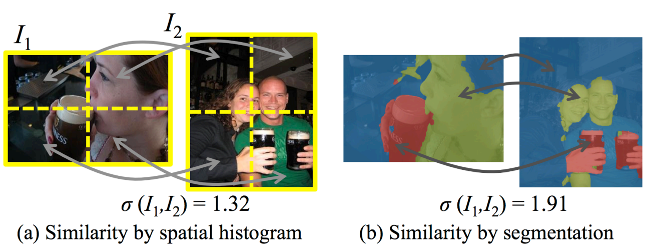

Figure 4: The benefit of segmentation for image similarity measurement. (a) For an unsegmented image pair, the spatial pyramid histograms are constructed on the whole images, which may not correctly reflect the location and scale variations. (b) After segmentation, the image similarity is computed as the mean of the best assigned segment similarities.

image pair, the image similarity is calculated from two-level spatial pyramid histograms on the whole images [17], which are not robust against location, scale, and pose variation as shown in Fig.4.(a). On the other hand, as shown in Fig.4.(b), this issue can be largely alleviated even with an imperfect segmentation. Given the two sets of segments of the images, we find the best matches between them by solving the linear assignment problem. Then, we compute the image similarity by the mean of similarities between matched segments. For the segment similarity, we use the histogram intersection kernel on the spatial pyramids of the segments.

3.3 Constructing K-Nearest Neighbor Graph

Given the image descriptors and similarity measures, the construction of a KNN graph is straightforward. However, if we naively compare all pairwise similarity by brute-force, it takes $\mathcal{O}(N^2)$, which can be prohibitively slow for a large $\mathcal{I}$. Fortunately, a large number of algorithms have been developed to construct approximate KNN graphs with avoiding the quadratic complexity. In this paper, we use the idea of multiple random divide-and-conquer [27], which allows to create an approximate KNN graph of high accuracy in $\mathcal{O}(N \log N)$ time. The method is simple: we randomly and recursively partition the dataset into subsets, and build an exact neighborhood graph over each subset. This random divide-and-conquer process repeats for several times, and then the aggregation of all neighborhood graphs of subsets can create a more accurate approximate KNN graph with a high probability. The details of procedures and theoretic analyses can be found in [27]. In our experiments, meta-data of images are also exploited for recursive random division. We repeat partitioning the image set into subsets according to each type of meta-data (*e.g.* image sources, owners, titles, or timestamps, if available). For example, in one partition, each subset includes the images taken at similar time; in another partition, each subset comprises the images owned by the same user. The basic assumption is that if images are taken at similar time or by the same user, they are likely to share similar contents. In our experiments, this heuristics of meta-data is efficient and effective to build KNN graphs.

3.4 Exemplar detection and clustering

Given a KNN graph $\mathcal{G}$, our next step is to perform exemplar detection. As a base algorithm, we use the diversity ranking algorithm of [15], which can choose L number of exemplars that are not only most central but also distinctive one another, by solving submodular optimization on the similarity graph $\mathcal{G}$. Since the L exemplars are discovered in a decreasing order of ranking scores, one can set L to an arbitrary large number. In this paper, we do not discuss the details of the algorithm, which can be found in [15]. Instead, we denote the exemplar detection procedure by $\mathcal{A} = \texttt{SubmDiv}(\mathbf{G}, L)$ where $\mathcal{A}$ is the set of exemplars and

Algorithm 1: Exemplar detection and clustering.

Input: (1) Image graph $\mathbf{G}$. (2) Number of exemplars L.
Output: (1) Exemplar set $\mathcal{A}$ and cluster set $\mathcal{C}$.

1: Append a constant vector $\mathbf{z} \in \mathbb{R}^{(N+1)\times 1}$ to the end column of $\mathbf{G}$ and $\mathbf{z}^T$ to the end row of $\mathbf{G}$. ($N = |\mathbf{G}|$).
2: $\mathcal{A} = \texttt{SubmDiv}(\mathbf{G},\ M)$.
3: $\{\mathcal{C}_l\}_{l=1}^{L} = \texttt{ClustSrc}(\mathbf{G},\ \mathcal{A})$.

/* Select M number of central and diverse exemplars $\mathcal{A}$.
Function $[\mathcal{A}] = \texttt{SubmDiv}(\mathbf{G},\ M)$
- **1**: $\mathcal{A} \leftarrow \emptyset$. $\mathbf{u} = \mathbf{0} \in \mathbb{R}^{N\times 1}$.
- **while** $|\mathcal{A}| \le L$ **do**
  - **2**: **for** $i = 1 : N$ **do** $\mathbf{u}(i) = \texttt{TempSrc}(\mathbf{G}, \{\mathcal{A} \cup i\})$.
  - **3**: $\mathcal{A} \leftarrow \mathcal{A} \cup \operatorname{argmax}_i \mathbf{u}$. Set $\mathbf{u} = \mathbf{0}$.

/* Get marginal gain u from the $\mathbf{G}$ and the node set $\mathcal{P}$.
Function $[u] = \texttt{TempSrc}(\mathbf{G},\ \mathcal{P})$
- **1**: Solve $\mathbf{u} = \mathbf{L}\mathbf{u}$ where $\mathbf{L}$ is the Laplacian of $\mathbf{G}$ under constraints of $\mathbf{u}(\mathcal{P}) = 1$ and $\mathbf{u}(N+1) = 0$.
- **2**: Compute the marginal gain $u = |\mathbf{u}|_1$.

/* Get cluster set $\mathcal{C}$ from the graph $\mathbf{G}$ and exemplars $\mathcal{A}$.
Function $\mathcal{C} = \texttt{ClustSrc}(\mathbf{G},\ \mathcal{A})$
- **1**: Let $L = |\mathcal{A}|$ and $L = |\mathbf{G}|$. $\mathcal{V}$ is vertext set of $\mathbf{G}$.
- **2**: Compute the matrix $\mathbf{X} \in \mathbb{R}^{(L-L)\times L}$ by solving $\mathbf{L}_u\mathbf{X} = -\mathbf{B}^T\mathbf{I}_s$ where if we let $\mathcal{X} = \mathcal{V}\backslash\mathcal{A}$, $\mathbf{L}_u = \mathbf{L}(\mathcal{X}, \mathcal{X})$, $\mathbf{B} = \mathbf{L}(\mathcal{A}, \mathcal{X})$, and $\mathbf{I}_s$ is an $L \times L$ identity matrix.
- **3**: Each vertex $v{\in}\mathcal{V}$ is clustered $c_v{=}\operatorname{argmax}_k \mathbf{X}(j,k)$.

$\mathbf{G} \in \mathbb{R}^{N\times N}$ is the adjacency matrix of the graph $\mathcal{G}$. The pseudocode is summarized in the step **1**–**2** of Algorithm 1.

Next, the clustering is performed using the random walk model [10]; each image i is associated with the exemplar that a random walker starting at i is most likely to reach first. Then, we cluster the images that share the same exemplar as the most probable destination. This procedure is implemented as a function `ClustSrc` of Algorithm 1.

3.5 Brand Localization via Cosegmentation

As the clustering output, we obtain the groups of coherent images $\mathcal{C} = \{\mathcal{C}_l\}_{l=1}^{L}$. The brand localization is achieved by separately applying the cosegmentation algorithm to each cluster. This separate cosegmentation scheme is more beneficial not only for parallelization but also for performance. Especially, for performance, it prevents cosegmenting the images of no commonality, which contradicts the basic assumption of cosegmentation algorithms. For instance, given the *Prada* brand, cosegmenting bag and jewelry images could be worsen than individually segmenting each image.

The goal of cosegmentation is to partition each image into foreground (*i.e.* the regions recurring across the images like *bags* in Fig.3.(d)) and background (*i.e.* the other regions). We select the MFC method [14] as our base cosegmentation algorithm, since it is scalable and has been successfully tested with Flickr user images. The MFC algorithm consists of two procedures, which are *foreground modeling* and *region assignment*. The foreground modeling step learns the appearance models for foreground and background, which are accomplished by using any region classifiers or their combinations. We use the Gaussian mixture model (GMM) on the RGB color space. The foreground models compute the values of any given regions with respect to the foregrounds and background, based on which the region assignment allocates the regions of an image via a combinatorial-auction style optimization to maximize the overall allocation values. More details of the algorithm can be referred to [14].

Algorithm 2: Brand localization via cosegmentation.

Input: (1) Cluster set $\mathcal{C} = \{\mathcal{C}_l\}_{l=1}^{L}$. (2) Image graph $\mathbf{G}$.
Output: (1) Set of segmented images $\mathcal{F}$ for each $i \in \mathcal{I}$.

foreach $\mathcal{C}_l \in \mathcal{C}$ **do**
- **1**: Find central image $c = \texttt{SubmDiv}(\mathbf{G}_l, 1)$ where $\mathbf{G}_l = \mathbf{G}(\mathcal{C}_l)$ is the subgraph of $\mathcal{C}_l$.
- **2**: Apply the unsupervised MFC algorithm [14] to $\{c \cup \mathcal{N}_c\}$ where $\mathcal{N}_c$ is the neighbor of c in the graph $\mathbf{G}_l$. As a result, we obtain segmented images $\mathcal{F}_{c\cup\mathcal{N}_c}$.
- **3**: Let $\mathcal{U}_l \leftarrow \mathcal{C}_l\backslash\{c \cup \mathcal{N}_c\}$. $\mathcal{F} \leftarrow \mathcal{F}_{c\cup\mathcal{N}_c}$.
- **while** $\mathcal{U}_l \neq \emptyset$ **do**
  - **4**: Sample an image i from $\{\mathcal{U}_l \cap \mathcal{N}_{\mathcal{F}}\}$.
  - **5**: Get foreground model $\{v_i\} = \texttt{FM}\ (\{\mathcal{N}_i \cap \mathcal{F}\})$.
  - **6**: Segment the image $\mathcal{F}_i = \texttt{RA}\ (i, \{v_i\})$.
  - **7**: $\mathcal{U}_l \leftarrow \mathcal{U}_l\backslash i$. $\mathcal{F} \leftarrow \mathcal{F} \cup \mathcal{F}_i$.

/* $\{v_i\} = \texttt{FM}\ (\mathcal{F}_i)$ is the function to learn foreground model $\{v_i\}$ of MFC [14] from the segmented images $\mathcal{F}_i$.
/* $\mathcal{F}_i = \texttt{RA}\ (i, \{v_i\})$ is the function to run region assignment of MFC [14] on image i using $\{v_i\}$.

For each cluster $\mathcal{C}_l$, we perform the cosegmentation by iteratively applying the foreground modeling and region assignment steps under the guidance of the subgraph $\mathcal{G}(\mathcal{C}_l)$ whose vertex set is $\mathcal{C}_l$. Its basic idea is that the neighboring images in $\mathcal{G}(\mathcal{C}_l)$ are visually similar, and thus they are likely to share enough commonality to be segmented together. Therefore, we iteratively segment each image i by using the learned foreground models from its neighbors in the graph. Then, the segmented image i is subsequently used to learn the foreground models for its neighbors' segmentation. That is, we iteratively run foreground modeling and region assignment by following the edges of $\mathcal{G}(\mathcal{C}_l)$. The overall algorithm is summarized in Algorithm 2. For initialization, as shown in step **1**–**2** of Algorithm 2, we run the unsupervised version of the MFC algorithm to the exemplar of C_l and its neighbors, from which the iterative cosegmentation starts.

4. BRAND ASSOCIATION MAPS

We visualize the clusters (or exemplars) in a circular layout in order to concisely represent both short-range and long-range interactions between them. We place the visual clusters by using two different metrics, the *radial distance* and *angular distance*, inspired by the Nielsen's method [2]:

1. The *radial distance* of a cluster reflects how strongly it associates with the brand. A larger cluster appears closer to the center of the map.
2. The *angular distance* between a cluster pair shows their closeness. The smaller the angular distance between the two is, the higher the correlation is.

Since Nielsen's mapping algorithm is unknown and no photo-based brand association mapping has been developed yet, we design a new embedding algorithm that satisfies the above requirements. Our objective is to calculate $(\mathbf{r}, \boldsymbol{\theta}) \in \mathbb{R}^{L\times 2}$, which are the polar coordinates of all clusters of $\mathcal{C}$. Algorithm 3 summarizes the whole mapping procedure.

Radial distances of clusters: According to the requirement 1, a larger cluster has a smaller radial distance (*i.e.*

Algorithm 3: Computing polar coordinates of clusters.

Input: (a) Cluster set $\mathcal{C} = \{\mathcal{C}_l\}_{l=1}^{L}$. (b) Image graph $\mathbf{G}$.
(c) Image sizes to be drawn $\mathbf{t} \in \mathbb{R}^{L\times 1}$.
Output: Polar coordinates $(\mathbf{r}, \boldsymbol{\theta}) \in \mathbb{R}^{L\times 2}$ of $\mathcal{C}$.

/* *Radial coordinates.* */
1: Compute transition matrix $\mathbf{P}$ by row-normalizing $\mathbf{G}$.
2: Solve Eq.(1) to get stationary distribution $\boldsymbol{\pi} \in \mathbb{R}^{N\times 1}$.
3: **foreach** $\mathcal{C}_a \in \mathcal{C}$ **do** compute $\pi_a = \sum_{i\in\mathcal{C}_a} \boldsymbol{\pi}(i)$.
4: Let $\pi_{min} = \min_{a\in\mathcal{C}} \boldsymbol{\pi}_a$ and $\pi_{max} = \max_{a\in\mathcal{C}} \boldsymbol{\pi}_a$.
5: **foreach** $\mathcal{C}_a \in \mathcal{C}$ **do** obtain $\mathbf{r}(a)$ by solving Eq.(2).
/* *Angular coordinates.* */
6: Obtain the cluster similarity $\mathbf{S} \in \mathbb{R}^{L\times L}$ from Eq.(4).
7: Initialize $\boldsymbol{\theta}$ by polar dendrogram of hierarchical clustering on $\mathbf{S}$, $J = 0, J_{old}$ = a large number.
while $|J - J_{old}| > \epsilon$ **do**
  8: Calculate $\frac{\partial}{\partial\boldsymbol{\theta}} J \in \mathbb{R}^{L\times 1}$. For each $a \in \mathcal{C}$, $\frac{\partial}{\partial\boldsymbol{\theta}_a} J = \sum_{b\in\mathcal{C}} \left(\mathbf{S}(a,b) - \gamma|\theta_a - \theta_b|^{\gamma-1}\right) G$ where $G = -2(1-\cos(\theta_a - \theta_b))^{-1/2}(-\sin\theta_a\cos\theta_b + \cos\theta_a\sin\theta_b)$.
  9: $\boldsymbol{\theta}_{new} = \boldsymbol{\theta} + \mu\frac{\partial}{\partial\boldsymbol{\theta}} J$.
  10: $J_{new} = \sum_a\sum_b \mathbf{S}(a,b)|\theta_a - \theta_b| - \sum\sum|\theta_a - \theta_b|^{\gamma}$.
  11: Update $J_{old} = J, J = J_{new}, \boldsymbol{\theta} = \boldsymbol{\theta}_{new}$.
/* *Force-directed refinement.* */
12: Obtain Cartesian coordinates $\mathbf{x} \in \mathbb{R}^{L\times 2}$ from $(\mathbf{r}, \boldsymbol{\theta})$ and a pariwse distance matrix $\mathbf{D}$. Store the original $\mathbf{x}_0$.
while $\mathbf{x}$ *is updated* **do**
  13: Set the displacement vector $\mathbf{d} = \mathbf{0}$. Set attractive and repulsive forces: $f_a(x) = x^2/k$ and $f_r(x) = k^2/x$.
  foreach *pair* (a, b) *if* $\mathbf{D}(a,b) < \gamma(\mathbf{t}(a) + \mathbf{t}(b))$ **do**
    14: $\mathbf{d}(b) += f_r(|\mathbf{x}(b) - \mathbf{x}(a)|)$.
  15: **foreach** $a \in \mathcal{C}$ **do** $\mathbf{d}(a) -= f_a(|\mathbf{x}(a) - \mathbf{x}_0(a)|)$.
  16: **foreach** $a \in \mathcal{C}$ **do** $\mathbf{x}(a) += \mathbf{d}(a)$.
17: Obtain the final $(\mathbf{r}, \boldsymbol{\theta})$ from $\mathbf{x}$.

closer to the center). In order to estimate the cluster sizes, we first compute the stationary distribution $\boldsymbol{\pi} \in \mathbb{R}^{N\times 1}$ of the graph $\mathcal{G}$, where $\boldsymbol{\pi}(i)$ indicates a random walker's visiting probability of node i. We assume that the size of cluster $\mathcal{C}_a$ is proportional to the sum of stationary distribution of the nodes in $\mathcal{C}_a$, which means the portion of time that a random walker traversing the graph stays in the cluster $\mathcal{C}_a$. That is, in a larger cluster, a random walker stays longer.

Given the transition matrix $\mathbf{P}$ obtained by normalizing the rows of $\mathbf{G}$, the stationary probability vector $\boldsymbol{\pi}$ can be computed by solving $\boldsymbol{\pi} = \mathbf{P}^T\boldsymbol{\pi}$ with $\|\boldsymbol{\pi}\|_1 = 1$. However, it is well known from the success of PageRank that a regularized stationary distribution is more robust and can incorporate a prior knowledge; it can be obtained by solving

$$\boldsymbol{\pi} = \widetilde{\mathbf{P}}^T\boldsymbol{\pi} \quad \text{where } \widetilde{\mathbf{P}} = \lambda\mathbf{P} + (1-\lambda)\mathbf{1}\mathbf{v}^T \tag{1}$$

where $\mathbf{v} \in \mathbb{R}^{N\times 1}$ is the teleporting probability such that $v_i \geq 0, \|\mathbf{v}\|_1 = 1$. It can supply a prior ranking to each node; without it, one can let $\mathbf{v} = [1/N, \cdots, 1/N]^T$ be uniform. $\mathbf{1}$ is an all-one vector, and λ is a regularization parameter to weight the random walker's behavior between edge following and random transporting. We set $\lambda = 0.9$ in all experiments.

Once we have $\boldsymbol{\pi}$, then we compute the stationary probability π_a of each cluster $\mathcal{C}_a$ by summing over the values of vertices in the cluster: $\pi_a = \sum_{i\in\mathcal{C}_a}\boldsymbol{\pi}(i)$. Let r_{max} and r_{min} be max and min radius of the circular layout, and π_{max} and π_{min} be max and min cluster stationary probability, respectively. Finally, the radial coordinate $\mathbf{r}(c)$ of cluster $\mathcal{C}_a$ is

$$\mathbf{r}(a) = \frac{r_{max} - r_{min}}{\pi_{max} - \pi_{min}}(\pi_{max} - \pi_a) + r_{min}. \tag{2}$$

Angular coordinates of clusters: In order to obtain the angular coordinates $\boldsymbol{\theta}$ of clusters $\mathcal{C}$, we first compute all pairwise similarities $\mathbf{S} \in \mathbb{R}^{L\times L}$ between the clusters, and then apply the modified spherical *Laplacian Eigenmap* technique [3, 4] to project the clusters on a circular manifold.

We use the *random walk with restart* (RWR) algorithm [25] to define the cluster similarity on a graph. The similarity values of all nodes $\mathbf{s}_a$ with respect to cluster $\mathcal{C}_a$ is defined as

$$\mathbf{s}_a = \lambda\mathbf{P}\mathbf{s}_a + (1-\lambda)\mathbf{v}_a^T \text{with } \mathbf{v}_a(i) = \begin{cases} 1/|\mathcal{C}_a| & \text{if } i \in \mathcal{C}_a \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

The score $\mathbf{s}_a(i)$ means the probability that a random walker stays at node i when the walker follows the edge of graph with probability λ and return to uniformly random nodes of cluster $\mathcal{C}_a$ with $1-\lambda$. It is straightforward to compute the similarity score from $\mathcal{C}_a$ to $\mathcal{C}_b$, denoted by $\mathbf{S}(a,b)$, as follows:

$$\mathbf{S}(a,b) = \sum_{i\in\mathcal{C}_b}\mathbf{s}_a(i)/S_a \quad \text{where } S_a = 1 - \sum_{i\in\mathcal{C}_a}\mathbf{s}_a(i). \tag{4}$$

Next, we project the clusters on a unit circle from the $\mathbf{S}$. Our circular embedding is based on the *Spherical Laplacian Information Maps* (SLIM) [4], which extends the Laplacian eigenmap (LEM) optimization [3] with an additional constraint of embedding data on the surface of sphere.

Conceptually, if a pair of clusters is similar to each other, then their angular difference in embedding should be small. Hence, the objective is formulated as finding $\boldsymbol{\theta}$ to minimize

$$\boldsymbol{\theta} = \text{argmin}\sum_{a\in\mathcal{C}}\sum_{b\in\mathcal{C}}\mathbf{S}(a,b)|\theta_a - \theta_b| - \sum_{a\in\mathcal{C}}\sum_{b\in\mathcal{C}}|\theta_a - \theta_b|^{\gamma}. \tag{5}$$

The LEM objective (*i.e.* the first term of Eq.(5)) enforces nearby points in the graph to be as close together as possible in the angular representation. However, the optimization using only the LEM objective attains a trivial solution to collapse all data to the same point. Therefore, the regularizer (*i.e.* the second term) is included in order to spread the embedded clusters on a circle. It leads the optimization to prefer large angular distances between all pairs of clusters. We set the constant $\gamma = 0.5$ in our experiments.

Since the optimization problem in Eq.(5) has no closed-form solution, we employ a gradient descent procedure, as summarized in step **7**–**11** of Algorithm 3. By nature, the final embedding highly depends on the initialization, for which we first perform hierarchical clustering on the $\mathbf{S}$, and then use its polar dendrogram. This initialization enables similar nodes to have small geodesic distances.

Layout refinement: We slightly update the coordinates of clusters $(\mathbf{r}, \boldsymbol{\theta})$ so that the final visualization is more visually pleasant. More specifically, we separate any pair of exemplars that are too much overlapped, by using Fruchterman and Reingold's method, one of popular force-directed drawing algorithms. The cluster positions are updated to reach equilibrium states by the attractive and repulsive forces. The attractive forces encourage the updated positions to be as similar to the original $(\mathbf{r}, \boldsymbol{\theta})$ as possible, while the repulsive forces take part severely overlapped exemplars. This refinement is summarized in step **12**–**17** of Algorithm 3.

5. EXPERIMENTS

In our experiments, we first present the brand association maps for several competing brands in section 5.1. Then, we quantitatively evaluate the proposed approach from two technical perspectives: exemplar detection/clustering in section 5.2, and brand localization via image cosegmentation in section 5.3. Since the main goal of this paper is to achieve the two technical visualization tasks for brand associations, we focus on validating the algorithmic performance over other candidate methods instead of user study. Finally, we examine the correlation between our findings from community photos and the actual sales data of brands in section 5.4. We plan to make public our MATLAB code.

5.1 Visualization of Brand Association Maps

We present the brand association maps of six competing brands of the *luxury* category in Fig.5. We show top 20 exemplars (*i.e.* cluster centers) in the map. We make several interesting observations as follows. First of all, our algorithm successfully discovers brands' characteristic visual themes (*e.g.* watch clusters in the *Rolex* and the iconic check patterns of the *Burberry*). Second, much of highly ranked clusters attribute to some specific scenes where photo-taking is preferred. For example, in the *Rolex*, the clusters of horse-riding and auto-racing events that are sponsored by the *Rolex* are as dominant as those of its main product watches. Such event topics are more favorable to be recorded as pictures rather than texts. In the *Louis+Vuitton*, there are lots of *wedding* related clusters, which makes sense because the wedding is not only an event where the products of luxury brands are purchased much, but also a memorable moment where the photos are taken a lot.

Although our photo-based brand association map is novel and promising, there are several issues to be explored further. First, we may need to correctly handle highly redundant or noisy clusters, which are mainly caused by the imperfection of image processing and clustering. Second, we also need to deal with polysemous brand names; for example, the *Mont+Blanc* is also the name of the mountain, and the *Corona* indicates the astronomical phenomenon as well. If we use additional keywords during image crawling to filter them out, the volume of retrieved images decreases severely.

5.2 Results on Clustering

Task: We evaluate the performance of our algorithm for the exemplar detection/clustering task, by comparing with several candidate methods. For quantitative evaluation, we first choose 20 brands (*i.e.* five brands per category), and generate 100 sets of groundtruth per brand as follows. We randomly sample three images (i, j, k) from the image set of a brand, and manually label which of j and k is more similar to image i. We denote $j \succ k|i$ if j is more similar to i than k. Although the labeled sets are relatively few compared to the dataset size, in practice this sampling-based annotation is commonly adopted in standard large-scale benchmark datasets such as ImageNet [7] and LabelMe [22].

After applying each algorithm, suppose that C_i, C_j, and C_k denote the clusters that include image i, j, and k, respectively. Then, we compute the similarity between clusters $\sigma(C_j, C_i)$ and $\sigma(C_k, C_i)$ by using the RWR algorithm in section 4. Finally, we compute the accuracy of the algorithm using the Wilcoxon–Mann–Whitney statistics:

$$ACC := \frac{\sum_{(i,j,k)} \mathbb{I}(j \succ k|i \ \wedge \ \sigma(C_j, C_i) > \sigma(C_k, C_i))}{\sum_{(i,j,k)} \mathbb{I}(j \succ k|i)} \quad (6)$$

where $\mathbb{I}$ is an indicator function. The accuracy increases only if the algorithm can partition the image set into coherent clusters, and the similarities between clusters coincide well with human's judgment on the image similarity.

Baselines: We compare our algorithm with four baselines. The (`KMean`) and the (`Spect`) are the two popular clustering methods, K-means and spectral clustering, respectively. The (`LP`) is a label propagation algorithm for community detection [20], and the (`AP`) is the *affinity propagation* [8], which is a message-passing based clustering algorithm. Our algorithm is tested in two different ways, according to whether image segmentation is in a loop or not. The (`Sub`) does not exploit the image cosegmentation output, whereas the (`Sub-M`) is our fully geared approach. That is, this comparison can justify the usefulness of our alternating approach between clustering and cosegmentation. We set $L = 300$, and use the same image features in section 3.1 for all the algorithms.

Quantitative results: Fig.6 reports the results of our algorithm and four baselines across 20 brand classes. The leftmost bar set is the average accuracies of 20 classes. In most brand classes, the accuracies of our method (`Sub-M`) are better than those of all the baselines. The average accuracy of our (`Sub-M`) is 62.0%, which is much higher than 51.7% of the best baseline (`AP`). In addition, the average accuracies of the (`Sub-M`) are notably better than (`Sub`), which implicates that the cosegmentation for brand localization can improve the clustering performance as expected.

5.3 Results on Brand Localization

Task: The brand localization task is evaluated as follows. As groundtruths, we manually annotate 50 randomly sampled images per brand, for the same 20 brands in the previous experiments. We do not label too obvious images depicting products on white background, since they cannot correctly measure the performances of algorithms. The accuracy is measured by the intersection-over-union metric $(GT_i \cap R_i)/(GT_i \cup R_i)$, where GT_i is the groundtruth of image i and R_i is the regions segmented by the algorithm. It is a standard metric in segmentation literature [14, 15]. We compute the average accuracy from all annotated images.

Baselines: We select two baselines that can discover and segment the regions of objects from a large number of images in an unsupervised manner (*i.e.* with no labeled seed images). The (`LDA`) [21] is an LDA-based unsupervised localization method, and the (`COS`) [15] is a state-of-art submodular optimization based cosegmentation algorithm. Our algorithm is tested in three different versions, according to whether exemplar detection/clustering is in a loop or not. The (`MFC`) runs our cosegmentation without involving our clustering output (but using a random partitioning instead), in order to show the importance of the clustering step when segmenting highly diverse Web images. The (`MFC-S`) is a single loop of our exemplar detection/clustering and cosegmentation, and (`MFC-M`) iterates this process more than twice. In almost all cases, it converges in two iterations. Hence, this comparison can quantify the accuracy increase by the iterative algorithm. We run all algorithms in an unsupervised way for a fair comparison. Since it is hard to know the best

Figure 5: Examples of brand association maps for six brands of the *luxury* category.

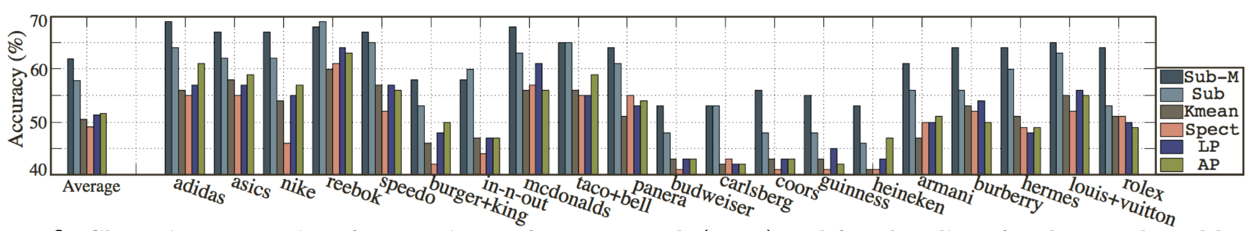

Figure 6: Clustering accuracies of two variants of our approach (`Sub-*`) and four baselines for the 20 selected brands. The average accuracies over the 20 brands, shown in the leftmost bar set, are (`Sub-M`): 62.0%, (`Sub`): 57.8%, (`Kmean`): 50.5%, (`Spect`): 49.2%, (`LP`): 51.4%, and (`AP`): 51.7%.

number of foregrounds K in advance (*e.g.* multiple foregrounds may exist in each image), we repeat each method by changing K from one to five, and report the best results.

Quantitative results: Fig.7 shows that our method outperforms other candidate methods in almost all classes. Especially, our average accuracy is 49.5%, which is notably higher than 36.7% of the best baseline (`COS`). In addition, the average accuracy of the (`MFC-M`) is also higher than those of (`MFC-S`) and (`MFC`), which demonstrates that the clustering and cosegmentation are mutually-rewarding.

Qualitative analysis: Fig.8 shows six sets of brand localization examples. The images of each set belong to the same cluster, and thus are cosegmented. We observe that the subjects of pictures and their appearances severely vary even though they are associated with the same brands. However, our approach is able to quickly cluster a large-scale image set and segment common regions in an unsupervised and bottom-up way, which can be an useful function for various Web applications, including detecting regions of brand for online multimedia advertisement.

5.4 Correlation with Sales Data

Since our work is the first attempt on exploring online photo collections for brand associations, we additionally report the statistics of correlations between image data and sales data of the brands. We conduct two different comparisons. First, we observe how the photo volumes of brands are correlated with their market shares. For example, the average annual revenue of the *Nike* is higher than that of the *Adidas* by about 40% from 2006 to 2011. We examine whether the *Nike* is also dominant over the *Adidas* in the volumes of Web photos. Second, we study in-depth correlation between the product groups of each brand. For example, the annual reports of the *Louis+Vuitton* classify their business into several product groups such as leather goods, perfume, jewelry, and wine. We compare between the proportions of product groups in image data and sales data of the brand.

We obtain the sales data from the annual reports that are publicly available on the companies' webpages. We ignore the brands held by private companies (*e.g. Chanel*), because it is often hard to know accurate financial information. In this analysis, we use image and sales data from 2006 to 2011.

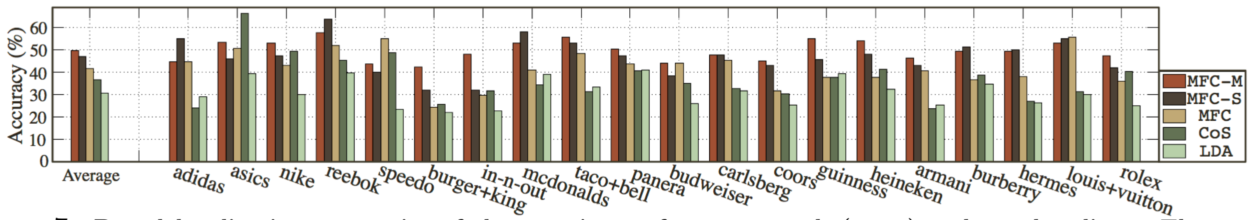

Figure 7: Brand localization accuracies of three variants of our approach (MFC-*) and two baselines. The average accuracies of the leftmost bar set are (MFC-M): 49.5%, (MFC-S): 46.8%, (MFC): 41.7%, (COS): 36.7%, and (LDA): 30.6%.

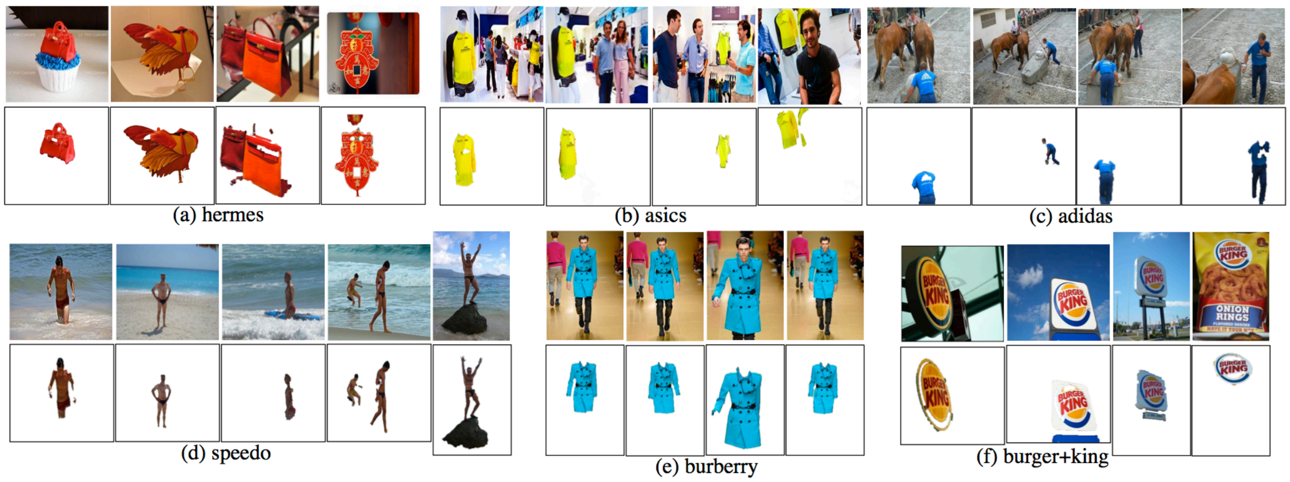

Figure 8: Six groups of brand localization examples. We sample four or five images per group that belong to the same cluster, and thus are jointly segmented. We show input images (top) and their segmentation output (bottom).

Correlation between photo volumes and market shares: Fig.9 shows the proportions of photo volumes and market shares for the brands per category. The ranking of the brands in the two data types are roughly similar, but the percentages do not agree each other because the preferred scenes or situations of photo taking are different from those of product purchase. For example, the *Guinness* has a larger percentage value in the photo volume than in the sales thanks to its positioning as premium beer. On the contrary, *Taco+Bell* occupies a small portion of photo volumes. It may be because the *Taco+Bell* is a cheap fastfood brand, which hardly attracts people to take pictures for the brand.

Correlation between product groups: Now we turn to the comparison between product groups in each brand. The main challenge here is that it is difficult for both human and computers to correctly classify millions of images into the predefined product groups. For human, the data size is too large to manually classify them. For computers, there is no classifier applicable to noisy Web images with high accuracies. Thus, we take advantage of our exemplar detection/clustering results. We manually classify each exemplar into one of predefined groups, and all the images in the same cluster are labeled as the same. The classification of product groups is based on the brand's annual reports.

Fig.10 shows the results of product group analysis for four luxury brands. We first label exemplar images by one of three groups: *product*, *company*, and *personal*. The *product* group comprises the photos whose main contents are the products of the brand. The *company* group includes the images that are directly relevant to the brand not to any particular products. It consists of four subgroups: *advertisement*, *logo*, *shop*, and *event*. The final one is the *personal* group for the private pictures whose contents are not explicitly associated with brands.

We summarize several observations as follows. First, in most brands, the *personal* group is the first or second largest one, which may result from that people usually take pictures on personal matters. Second, the *company* group is also very popular; for examples, people are interested in luxurious *Louis+Vuitton*'s stores as much as its products. Moreover, the events hosted by brands are also popularly taken such as fashion shows, music concerts, and sports activities. Third, in the *product* group, one or two leading product types take the majority of photo volumes while some product segments like *wines*, *perfume*, and *jewelry* rarely appear.

6. CONCLUSION

In this paper, we addressed a problem of visualizing the brand associations by leveraging large-scale online photo collections. We developed a novel approach to jointly performing exemplar detection/clustering and brand localization in a mutually-rewarding way. With the experiments of about five millions of images for 48 brands, we have shown the superiority of our approach for the two visualization tasks over other candidate methods. The empirical results assured that our method can be a fundamental component to achieve our ultimate goal: inferring in-depth brand associations from Web images, which is a next direction of our future work.

7. REFERENCES

[1] D. A. Aaker. Measuring Brand Equity Across Products and Markets. *Cal. Manag. Rev.*, 38(3):102–120, 1996.

[2] N. Akiva, E. Greitzer, Y. Krichman, and J. Schler. Mining and Visualizing Online Web Content Using BAM: Brand

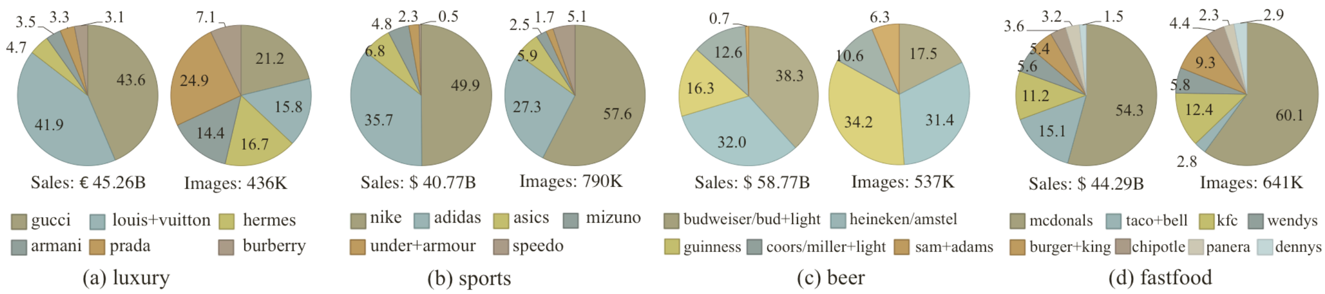

Figure 9: Comparison between the market shares (left) and the portions of photo volumes (right) for the brands of four categories: (a) *luxury*, (b) *sports*, (c) *beer*, and (d) *fastfood*. The numbers indicate percentage values.

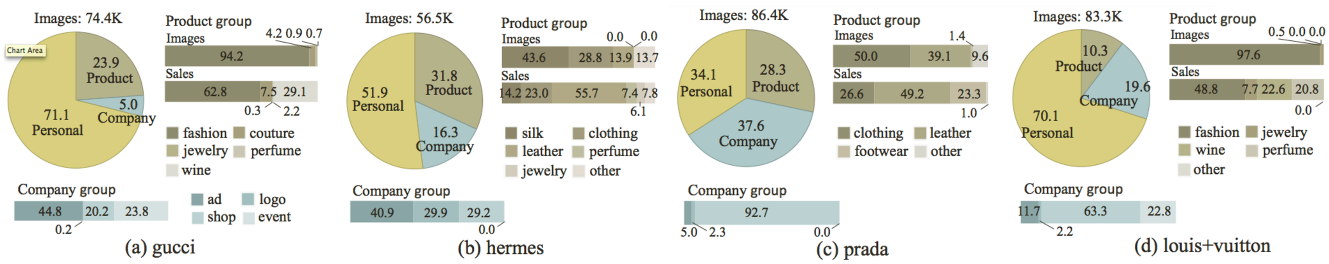

Figure 10: Results of the product group analysis for four luxury brands. Each pie chart shows the proportions of three groups in the image volume: *product*, *company*, and *personal*. In the bottom, the images of the *company* group are further classified into one of *advertisement*, *logo*, *shop*, and *event*. In the right, bar charts show the proportions of the images (top) and the actual revenues (bottom) for the *product* group. The classification of product groups is based on the brand's annual reports. The numbers indicate percentage values.

Association Map. In *ICWSM*, 2008.

[3] M. Belkin and P. Niyogi. Laplacian Eigenmaps for Dimensionality Reduction and Data Representation. *neural computation*, 15(6):1373–1396, 2003.

[4] K. M. Carter, R. Raich, and A. O. Hero. Spherical laplacian information maps (slim) for dimensionality reduction. In *SSP*, 2009.

[5] A. C. H. Chen. Using Free Association to Examine the Relationship between the Characteristics of Brand Associations and Brand Equity. *J. Product Brand Management*, 10(7):439–451, 2001.

[6] J. E. Danes, J. S. Hess, J. W. Story, and J. L. York. Brand Image Associations for Large Virtual Groups. *Qualitative Market Research*, 13(3):309–323, 2010.

[7] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In *CVPR*, 2009.

[8] B. J. Frey and D. Dueck. Clustering by Passing Messages Between Data Points. *Science*, 315:972–976, 2007.

[9] K. Gao, S. Lin, Y. Zhang, S. Tang, and D. Zhang. Logo Detection Based on Spatial-Spectral Saliency and Partial Spatial Context. In *ICME*, 2009.

[10] L. Grady. Random Walks for Image Segmentation. *IEEE PAMI*, 28:1768–1783, 2006.

[11] Y. Jing and S. Baluja. PageRank for Product Image Search. In *WWW*, 2008.

[12] H. Kang, M. Hebert, A. A. Efros, and T. Kanade. Connecting Missing Links: Object Discovery from Sparse Observations Using 5 Million Product Images. In *ECCV*, 2012.

[13] K. L. Keller. Conceptualizing, Measuring, and Managing Customer-Based Brand Equity. *J. Marketing*, 57(1):1–22, 1993.

[14] G. Kim and E. P. Xing. On Multiple Foreground Cosegmentation. In *CVPR 2012*, 2012.

[15] G. Kim, E. P. Xing, L. Fei-Fei, and T. Kanade. Distributed Cosegmentation via Submodular Optimization on Anisotropic Diffusion. In *ICCV 2011*, 2011.

[16] J. Kleban, X. Xie, and W.-Y. Ma. Spatial Pyramid Mining for Logo Detection in Natural Scenes. In *ICME*, 2008.

[17] S. Lazebnik, C. Schmid, and J. Ponce. Beyond Bags of Features: Spatial Pyramid Matching for Recognizing Natural Scene Categories. In *CVPR*, 2006.

[18] D. L. Nelson, C. L. McEvoy, and T. A. Schreiber. The University of South Florida Free Association, Rhyme, and Word Fragment Norms. *Behavior Research Methods, Instruments, Computers*, 36(3):402–407, 2004.

[19] Nielsen Online. Brand Association Map, 2010.

[20] U. N. Raghavan, R. Albert, and S. Kumara. Near Linear Time Algorithm to Detect Community Structures in Large-Scale Networks. *Phys Rev E*, 76(036106), 2007.

[21] B. C. Russell, A. Efros, J. Sivic, W. T. Freeman, and A. Zisserman. Using multiple segmentations to discover objects and their extent in image collections. In *CVPR*, 2006.

[22] B. C. Russell, A. Torralba, K. P. Murphy, and W. T. Freeman. LabelMe: A Database and Web-based Tool for Image Annotation. *IJCV*, 77:157–173, 2008.

[23] S. Sanyal and S. H. Srinivasan. LogoSeeker: A System for Detecting and Matching Logos in Natural Images. In *ACM MM*, 2007.

[24] O. Schnittka, H. Sattler, and S. Zenker. Advanced Brand Concept Maps: A New Approach for Evaluating the Favorability of Brand Association Networks. *I. J. Research in Marketing*, 2012.

[25] J. Sun, H. Qu, D. Chakrabarti, and C. Faloutsos. Neighborhood Formation and Anomaly Detection in Bipartite Graphs. In *ICDM*, 2005.

[26] B. D. Till, D. Baack, and B. Waterman. Strategic Brand Association Maps: Developing Brand Insight. *J. Product Brand Management*, 20(2):92–100, 2011.

[27] J. Wang, J. Wang, G. Zeng, Z. Tu, R. Gan, and S. Li. Scalable k-NN Graph Construction for Visual Descriptors. In *CVPR*, 2012.

[28] K. Yamaguchi, M. H. Kiapour, L. E. Ortiz, and T. L. Berg. Parsing Clothing in Fashion Photographs. In *CVPR*, 2012.