Understanding and Classifying Image Tweets

Tao Chen[1], Dongyuan Lu[1], Min-Yen Kan[1], Peng Cui[2]
[1]School of Computing, National University of Singapore
[2]Department of Computer Science and Technology, Tsinghua University

Introduction

Image Tweets

- 56% of all posts in Sina Weibo
- Retweeted more often and survived longer than text-only posts
- Important forms of tweets while few works have been conducted

The aims of this study are

Understanding

- Image characteristics
- Image Tweets vs. Text Tweets

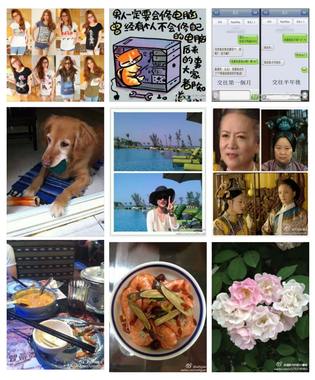
Classifying

- Image tweets by the image-text relations.

Dataset used in this study

- 57.6 million tweets from Weibo public timeline
- 45.1% are image tweets

Image Characteristics



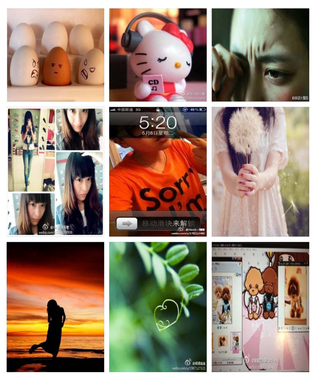
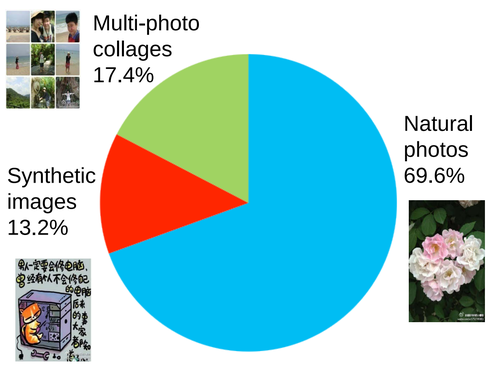
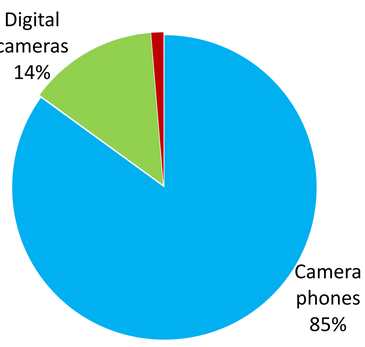


Image formats



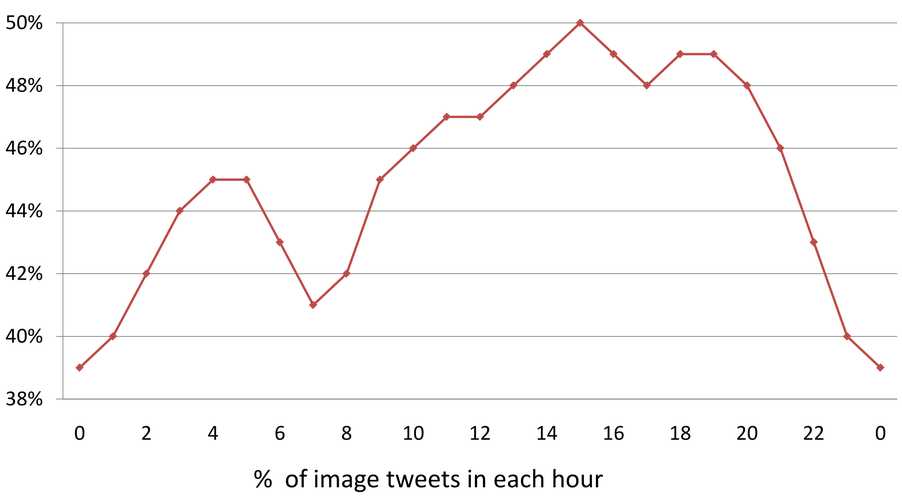
Devices used to capture photos



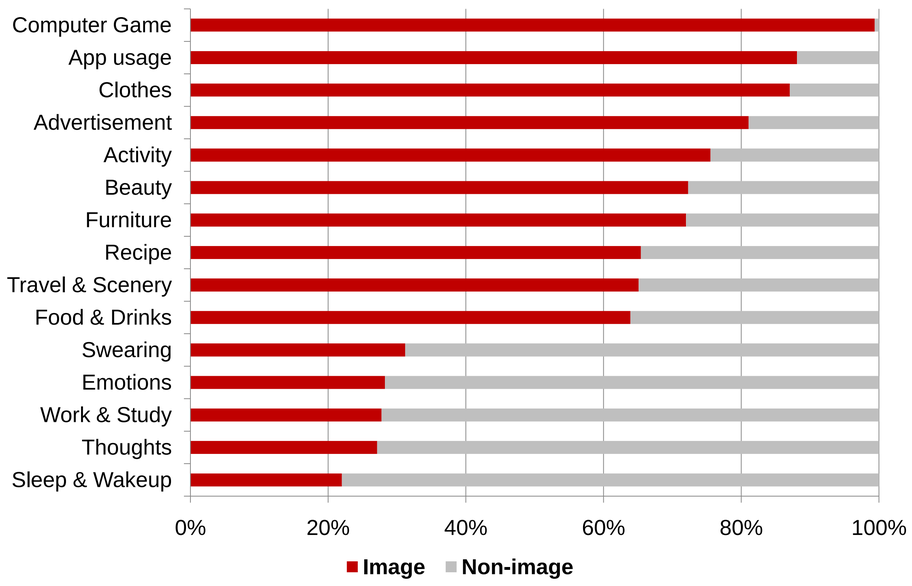
- Users care content more than photo quality

Image Tweets vs. Text Tweets

When



What



Why

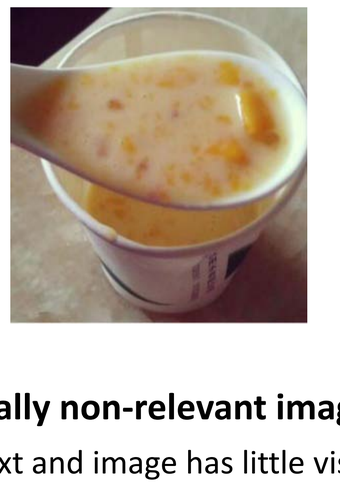
- Preference of posting image tweet or text tweet is highly correlated with the content

Image and Text Relations

Visually relevant image tweets (visual)

- At least one noun/verb in text corresponds to the whole or part of the image
- Text and image are complementary
  - Image: visual highlights
  - Text: contextual description

她们在厨房折腾了两个小时的结果……芒果椰汁西米露 (This is the result of 2 hours of their hard work in the kitchen ...Sago cream with mango and coconut)



陈建斌怎么看怎么还是曹操的样子啊! (No matter how I look at it, Chen Jianbing looks like Cao Cao!)

Visually non-relevant image tweets (non-visual)

- Text and image has little visual correspondence
  - Decorative images to attract readership
  - Emotional relevance

这次不知道又要隔几天才能见面了，我想你宝贝@FishSwing (It will be many days before I see you again, I miss you darling @FishSwing.)

可恶的蚊子，我要杀了你! (Horrible mosquitoes, I will kill you!)

Visual/Non-Visual Classification

Distinction of image-text relations is important for

- Text-based image retrieval
- Automated tagging generation
- Prioritizing image display in small screen

Therefore, we automate the distinction as a supervised binary classification

Gold dataset construction

- 4811 annotated image tweets: visual (66.6%) vs. non-visual (33.4%)
- Crowdscourced 72 annotators
- Each image was labeled by three annotators
- Fleiss' k: 0.62 (substantial agreement)
- Released at http://wing.comp.nus.edu.sg/downloads/imagetweets/

Method

- Naïve Bayes classifier (outperforms other classifiers)
- 10 fold cross-validation
- Best result: 70.5% at Macro-F_1

Class	Features	Macro-F_1 (%)
Text	(1) : Words only (Base line)	64.8
	(2) : (1) + Microblog-specific	65.2
	(3) : (1) + Named Entities	65.3
	(4) : (1) + Text Topic	66.6
	(5) : (1) + POS Density	69.7
Image	(6) : (1) + Image Topic	65.4
	(7) : (1) + Face	65.7
Context	(8) : (1) + Retweets	60.9 (-)
	(9) : (1) + Comments	64.5 (-)
	(10) : (1) + Replied by Author	64.7 (-)
	(11) : (1) + Device	64.9
	(12) : (1) + Follower Ratio	64.9
	(13) : (1) + Posting Time	65.0
All	(14) : (1-7 + 11-13)	70.5
Majority		40.0