

A HIERARCHICAL APPROACH TO PRACTICAL BEVERAGE PACKAGE RECOGNITION

Mei-Chen Yeh* and Jason Tai

Department of Computer Science and Information Engineering
 National Taiwan Normal University







Cost-conscious shopping: easy, fast, intuitive!

PRODUCT RECOGNITION



• Barcode, QR code





• Radio frequency identification (RFID)



Tags are required!

Content-based approach: recognize the product from any part of the content

Demo video



http://www.youtube.com/watch?v=FjZOHwaBL6Q

OUTLINE

- Introduction
- Challenges in beverage package recognition
- Approach
 - System framework
 - Recognition module
- Experiments
- Conclusion and future work

CHALLENGES

• Various package forms









CHALLENGES

• Various package forms • Arranged in any angle to users



Gog Cola

CHALLENGES

- Various package forms
- Arranged in any angle to users



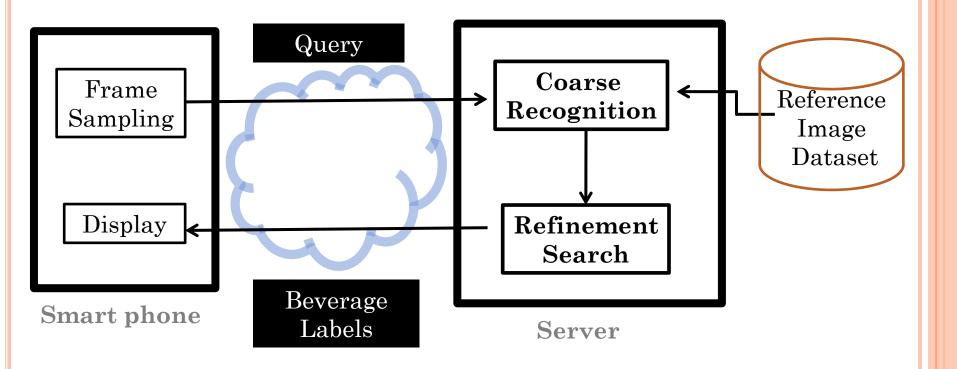
• Common visual patterns on different products



OUTLINE

- Introduction
- Challenges in beverage package recognition
- Approach
 - System framework
 - Recognition module
- Experiments
- Conclusion and future work

System Framework



REFERENCE IMAGE DATASET



- Around 100 samples (keeps growing)
- 73,026 interest points
- Panoramic images

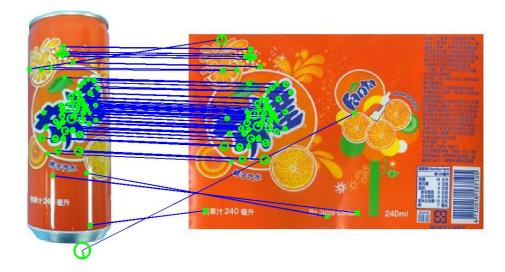
				0							
7-black-soymilk.jpg	7-english-tea-latte.jpg	apple_sidra.JPG	asparagus juice.JPG	black-water.jpg	bmilk1.jpg	bmilk2.jpg	bomi_mixed_juice.FG	bomi_today_juice.IPG	bomi-today-purple.jpg	bubble_greentea_10.FFG	bubble_greentea_20.JPG
cafeplaza.JPG	classic_blackmilktea.JPG	cocacola.JPG	coke_can.jpg	cranicetea.JPG	fanta_grape2.jpg	fanta_orange2.jpg	five_grape.jpg	five_lemon.jpg	five_orange.jpg	golden_soymilk.PG	green_milktea.FG
greenjelly.JPG	green-time-water.jpg	Passion Fruit Possion Fruit heysong_passionfruit.jpg	Apple beysong-apple.jpg	i_love_milk_low_fat.jpg	i-love-milk.jpg	jasmine_green_tea.jpg	FREE N Iemon_IU.JPG	lipton-green-milk-tea.jpg	mai_blacktea.JPG	mai_milktea.JPG	mangotea.JPG
matcha_milktea.jpg	moli_Gfruit_tea.jpg	moli_honeytea.jpg	mountpessionfruit.PG	nest_black_teal.jpg	oolong_milktea.JPG	original_milktea.JPG	ovaltine1.PG	ovaltine2.PG	papaya-milk.jpg	pessionfruit.JPG	pepsi.jpg
president-black-milk-te	president-green_milk_t	president-milk jpg	pure_coffe2.jpg	rebull2.jpg	refreshers_cran.PG	sars2.jpg	shin12tea.JPG	sprite.jpg	tiramisu jpg	tryit_blacktea.JPG	tryit_dist_greentea.jpg

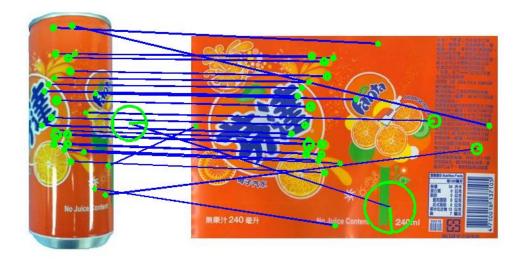
REFERENCE IMAGE DATASET

- Around 100 samples (keeps growing)
- 73,026 interest points
- Panoramic images









OBSERVATIONS

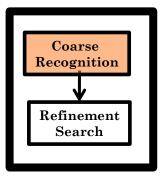
• Each local keypoint has different discriminative power

• Informative but not discriminative



• A keypoint's discriminative power may vary given different contexts





COARSE RECOGNITION

• Goals

- Filters out irrelevant images
- Determines the context for a refinement search
- Image representation
 - Local descriptors
 - Scale-invariant feature transform (SIFT), 128-d
 - Speeded up robust features (SURF), 64-d
- Image matching
 - Locality sensitive hashing (LSH)
 - Cosine similarity

Server

LSH: FAST MATCHING AND FEATURE ANALYSIS • Hash function: $h_{a,b} = \left| \frac{a \cdot v + b}{r} \right|$ $\mathbf{a} \cdot \mathbf{v} + b$

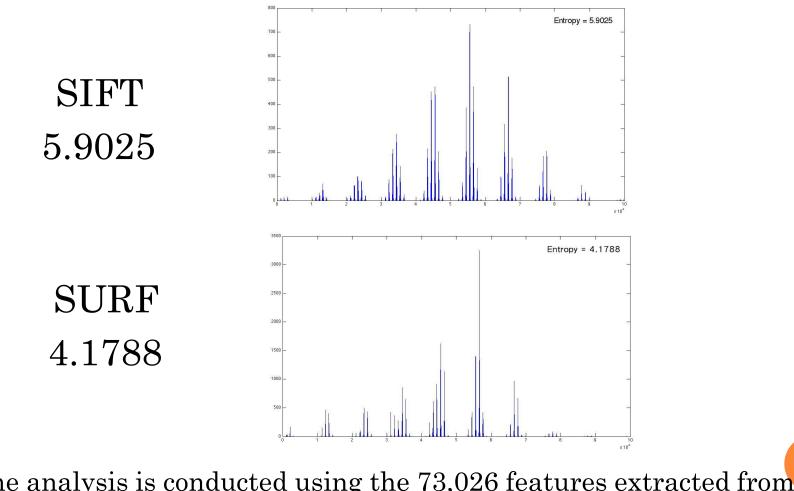
• Implementation: 5 random projections, 10 intervals, 10,000 buckets in total

Lee et al., "Probing the Local-Feature Space of Interest Points," ICIP 2010.

- SIFT, Berkeley natural images: 4.11
- SIFT, noise patches: 2.38

[Lee et al., ICIP 2010]

FEATURE ANALYSIS



The analysis is conducted using the 73,026 features extracted from ¹⁷ our reference image set.

REFINEMENT SEARCH

• Goals

- Selects the exact flavor in a beverage series
- Detects a new flavor that is not in our dataset
- Feature weighting function:

$$w(p_i) = \frac{N - \sum_k t_{i,k}}{N - 1}$$

 $p_i: i$ -th feature point, k: candidate index, N: number of candidates, $t_{i,k}:$ a binary variable that represents the presence/absence of a keypoint in the k-th candidate mage that matches p_i

• Example

- $w(p_i) = 1 : p_i$ appears exactly on one candidate image
- $w(p_i) = 0$: p_i appears on every candidate images



Coarse Recognition

Server

FEATURE WEIGHTING

Reference images



Query images







FEATURE WEIGHTING

- No need to construct a visual vocabulary
- Weights are not pre-computed
- Weights are query-dependent and adaptive

OUTLINE

- Introduction
- Challenges in beverage package recognition
- Approach
 - System framework
 - Recognition module
- Experiments
- Conclusion and future work

EXPERIMENTS (1)

• Test set:

 Select randomly 40 beverages, each has 3 images, 120 images in total



Exhaustive Search	LSH	Accuracy 92.5%	
73026 * 923 = 67,402,998	306,350	Sparseness 0.795%	
17.8905 (5.5385)	0.1289 (0.0368)	Speed-up 167.73x	
Exhaustive Search	LSH	Accuracy 76.67%	
72170 *1051 = 75,850,670	603,480	$\begin{array}{c} { m Sparseness} \\ 0.454\% \end{array}$	
11.0793 (2.3183)	0.1086 (0.0372)	Speed-up 150.50x	
	Search 73026 * 923 = 67,402,998 17.8905 (5.5385) Exhaustive 72170 *1051 = 75,850,670 11.0793	SearchLSH $73026 * 923 = \\ 67,402,998$ $306,350$ $17.8905 \\ (5.5385)$ $0.1289 \\ (0.0368)$ $17.8905 \\ (5.5385)$ $0.1289 \\ (0.0368)$ Exhaustive SearchLSH $72170 * 1051 = \\ 75,850,670$ $603,480$ 11.0793 0.1086	

FAILED CASES (9/120)



EXPERIMENTS (2)

• Test set:

- Beverage packages in a series
- Internet images
- Set 1: 24 images among 9 series (in our reference set, should be recognized)
- Set 2: 8 images (not in our reference set, should be rejected)





EXPERIMENTS (2)

• Set 1 (24): 22 images are recognized

• Set 2 (8): all images are rejected

Recognition accuracy: 91.66% (22/24)





CONCLUSION AND FUTURE WORK

- A practical beverage package recognition system
- The hierarchical approach
 - Calculates the query-dependent feature weights when needed
 - Performs fast as most images are filtered using a hashing scheme
- Future work
 - A new representation that incorporates color information
 - An indexing approach for visual instances described by multiple cues



THANK YOU

28

More information:

http://www.csie.ntnu.edu.tw/~myeh