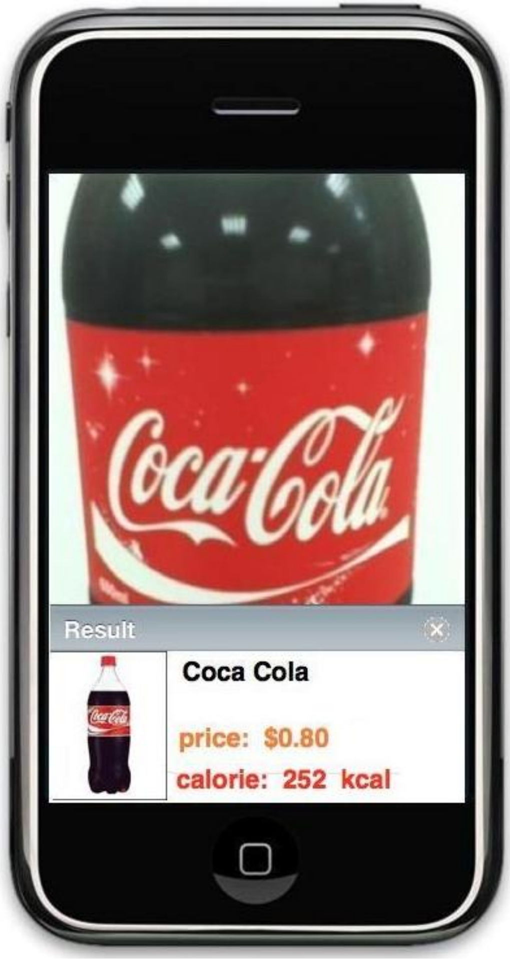
A Hierarchical Approach to Practical Beverage Package Recognition

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

. . .

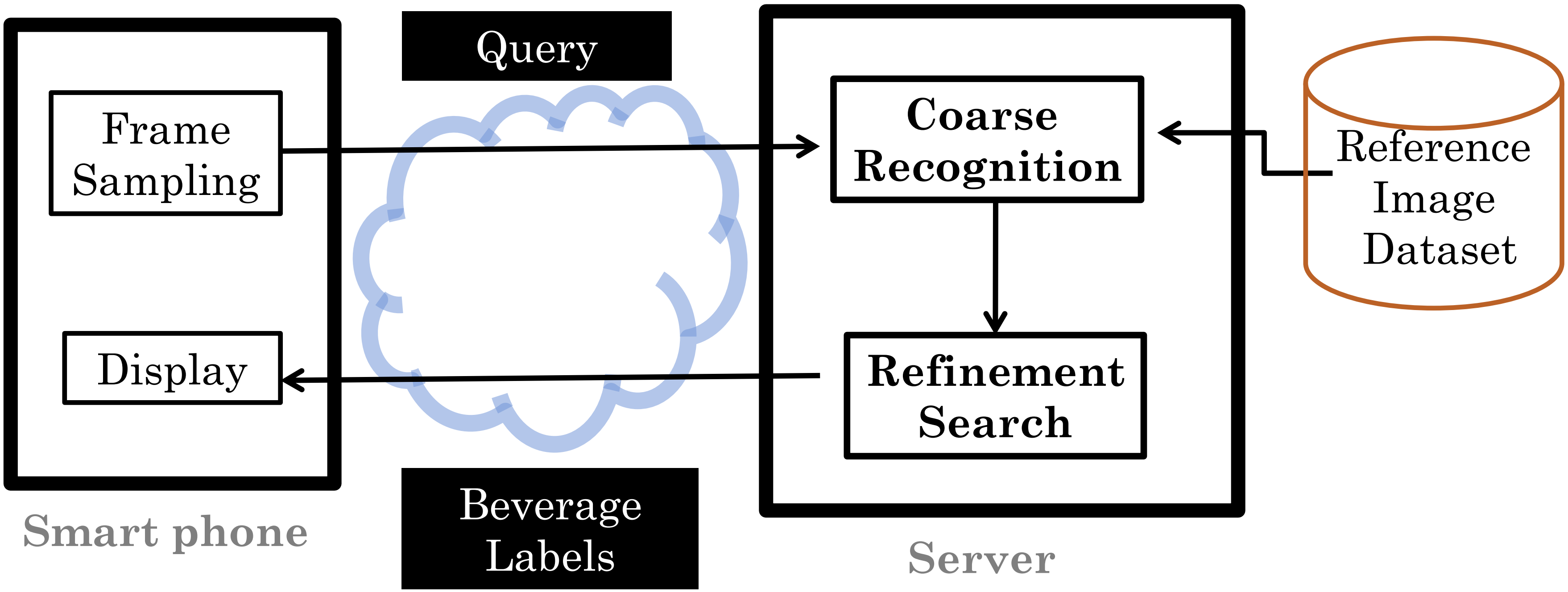
CHALLENGES

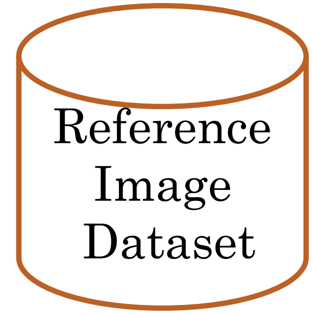
- Various package forms
- Arranged in any angle to users
- Common visual patterns on different products

Outline

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

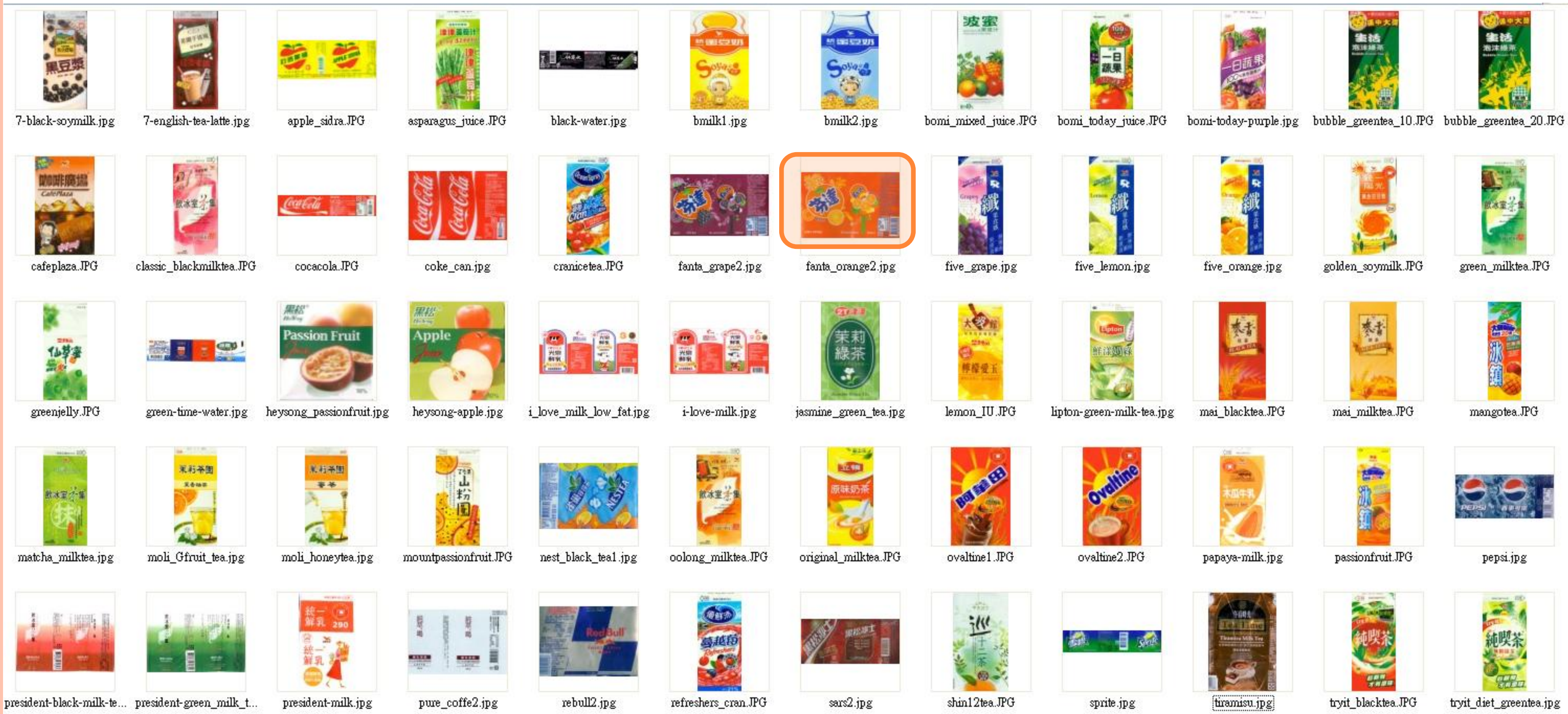
System Framework



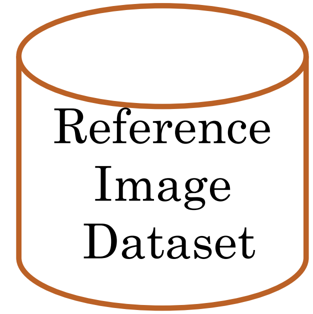


Reference Image Dataset

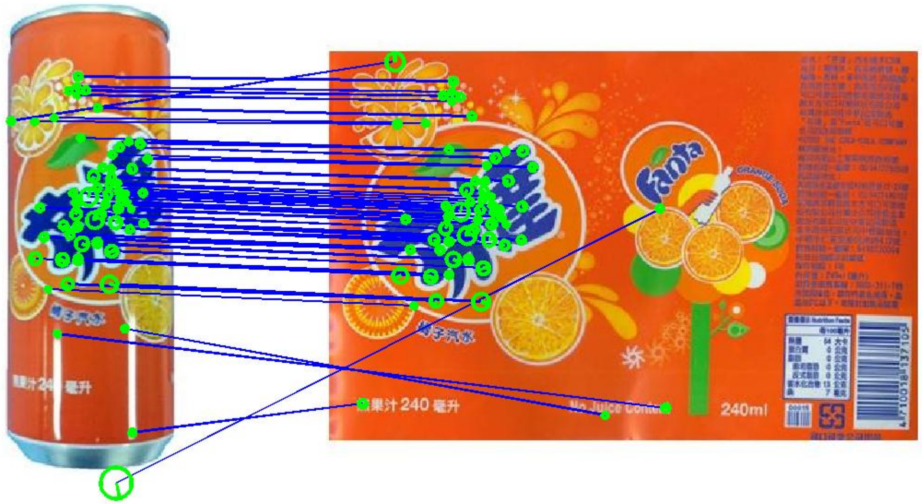
- Around 100 samples (keeps growing)
- 73,026 interest points
- Panoramic images

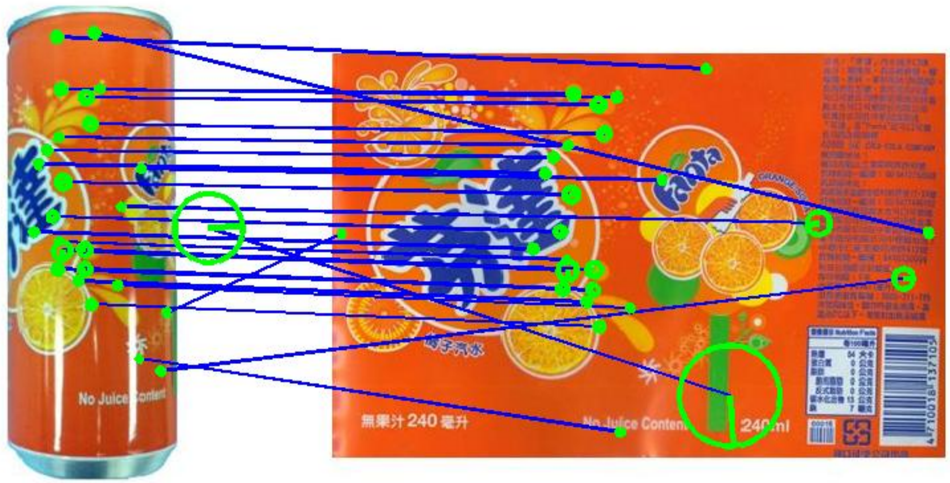


Reference Image Dataset



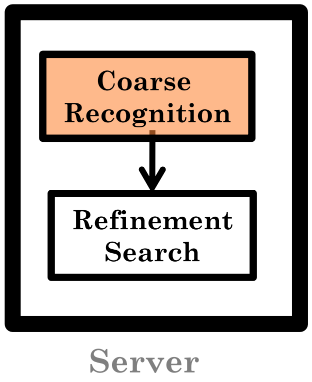
- Around 100 samples (keeps growing)
- 73,026 interest points
- Panoramic images


Fanta
無果汁 240 毫升
240ml


Fanta
No Juice Content
無果汁 240 毫升

Observations

- Each local keypoint has different discriminative power
  - Informative but not discriminative
- A keypoint's discriminative power may vary given different contexts

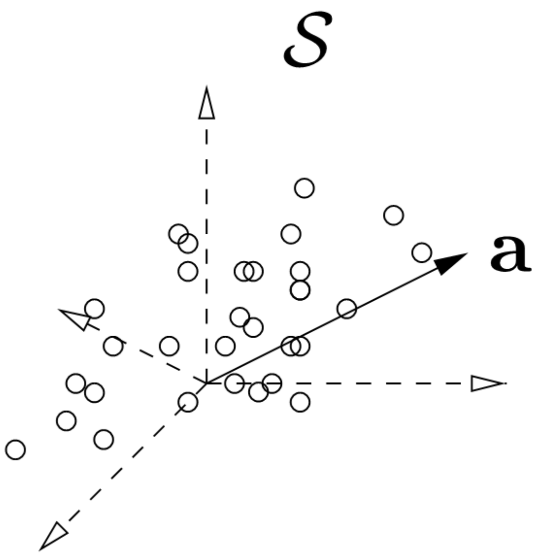


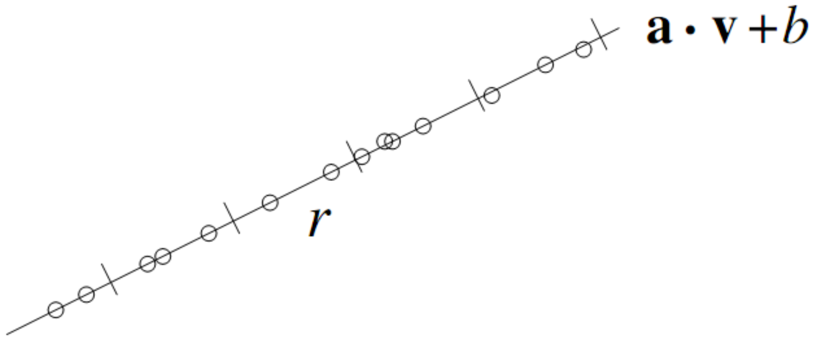
Coarse Recognition

- 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

LSH: Fast Matching and Feature Analysis

- Hash function: $h_{a,b} = \left\lfloor \frac{a \cdot v + b}{r} \right\rfloor$





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

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

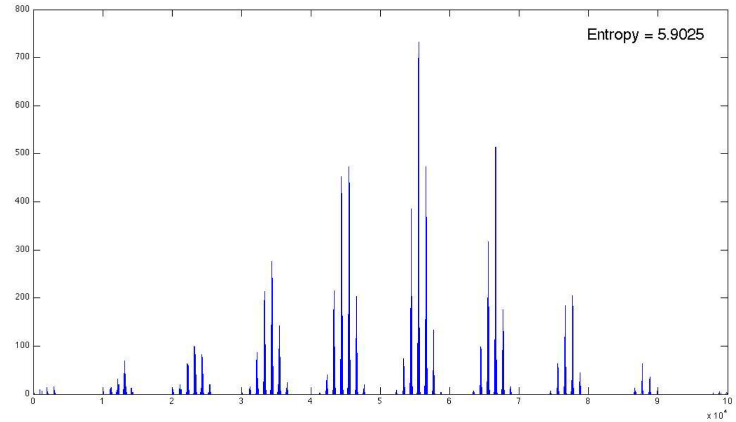
FEATURE ANALYSIS

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

[Lee *et al.*, *ICIP* 2010]

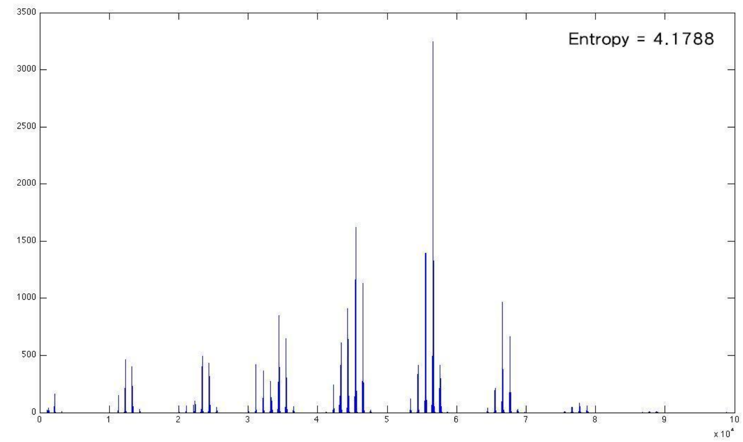
SIFT

5.9025

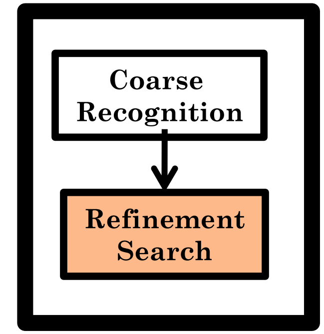


SURF

4.1788



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



Server

Refinement Search

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

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

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

- Example
  - $\boldsymbol{w(p_i) = 1}$: $\boldsymbol{p_i}$ appears exactly on one candidate image
  - $\boldsymbol{w(p_i) = 0}$: $\boldsymbol{p_i}$ appears on every candidate images

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

Box

Tetra Pak

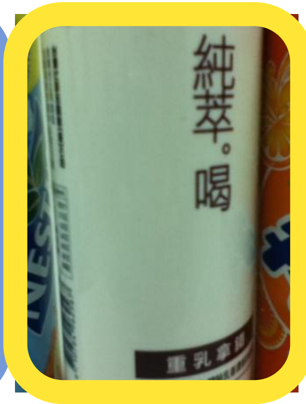
Can

Bottle

SIFT-128	Exhaustive Search	LSH	Accuracy 92.5%
Feature pairs of comparison	73026 * 923 = 67,402,998	306,350	Sparseness 0.795%
Time (seconds)	17.8905 (5.5385)	0.1289 (0.0368)	Speed-up 167.73x

SURF-64	Exhaustive Search	LSH	Accuracy 76.67%
Feature pairs of comparison	72170 *1051 = 75,850,670	603,480	Sparseness 0.454%
Time (seconds)	11.0793 (2.3183)	0.1086 (0.0372)	Speed-up 150.50x

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

More information:

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