CrowdSight
Rapidly Prototyping Intelligent Visual Processing Apps

Mario Rodriguez, James Davis
Computer Science - UCSC
My phone has a camera…

how can I leverage it to make my life easier?
Concept:

1. Take pic
2. Computation happens
3. Need is met
Great, I just need:

- Smart phone with camera
- Food calorie database
- Spreadsheet/charting software
- Computer vision
Minimum Viable Product

that version of a new product which allows a team to collect the maximum amount of validated learning about customers with the least effort
CrowdSight

- Framework for quickly prototyping mobile apps which require visual processing
- Leverages human computation to get work done
- Allows apps to be shared with others
- Targets smart and not-so-smart phones
Related Work

- Development tools for human computation
  - Turkit
  - QuickTurkit
  - hQuery

- Mobile apps using human computation
  - mCrowd
  - VizWiz
  - AmazonRemembers
System Architecture
Service Definition

title: Transcribe text in image
reward: 0.05
assignmentduration: 3600
qual.comparator: greaterthan
qual.value: 25

QUESTIONFORM

QUESTION

QUESTIONCONTENT

<Text>Transcribe text in image:</Text>

ANSWERSPECIFICATION

FREETEXTANSWER

<NumberofLines>10</NumberofLines>

</FREETEXTANSWER>

</ANSWERSPECIFICATION>

</QUESTION>

</QUESTIONFORM>
Image Workflow
Test Applications

- Counting people
- Product finder
- Business card scanning
Counting People

- Yang et. al. (2003) report several man-months to develop a real-time people counting system.

- With a US$70 Panasonic PetCam we had a system running up and running in less than 1 hour.

- US$0.01 per image @ 1 image per minute.
Product Finder

- Take a picture of something and get back a link to where the item can be purchased
- Ongoing research in object recognition with computer vision

Options:
- Amazon Remembers (limited to Amazon products)
- SnapTell (books, DVDs, CDs, video games)
- RedLaser (requires barcodes)

- 30 minutes to build using CrowdSight
Product Finder

- Tested with 35 products: tea, books, DVD’s, food, electronics, computer parts, furniture, etc
- Accuracy: is the link relevant and can it be purchased?

<table>
<thead>
<tr>
<th>System</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>SnapTell</td>
<td>10%</td>
</tr>
<tr>
<td>RedLaser</td>
<td>30%</td>
</tr>
<tr>
<td>CrowdSight</td>
<td>80%</td>
</tr>
</tbody>
</table>

- More complex flow:
  - Get several candidates
  - Vote on best/cheapest
Business Card Scanning

- Built in 15 minutes along with form with 12 fields for Turkers to fill out

- US$0.05 per card vs BizSnap US$1.00 per 10 cards

- CrowdSight: ~ 5 minutes to get a response and 80% accuracy (without redundancy)

- BizSnap (OCR): ~ 1 minute for response and 40% accuracy
Business Card Scanning

CardMunch
- ~ 5 minutes to get a transcription
- US$0.01 – US$0.10 cost per card (AMT)
- App allowed resubmits
- Acquired by LinkedIn in 2011

BizSnap
- < 1 min to get a response
- Can edit card after transcription
- US$1 for 10, US$2 for 25, US$6 for unlimited

VistaPrint in 2010:
- 52,000 orders/day
- US$250 million in sales

Acquired by LinkedIn in 2011
Future Work

- Aggregate existing services into composite workflows
- WYSIWYG interface for service definition
- More sophisticated quality control mechanisms
- Process additional kinds of media (audio, video)
- Open source code and/or host the service
Questions?