Mining and Summarizing Customer Reviews

MINQING HU AND BING LIU
DEPARTMENT OF CS, UIUC

PRESENTED BY
MD ANIDNYA PRODHAN
Motivation

Source: http://www.amazon.com/Canon-Rebel-Digital-18-55mm-Lens/dp/B00BW6LWO4
Motivation

Customer Reviews

🌟🌟🌟🌟🌟 (254)
4.7 out of 5 stars

- 5 star: 221
- 4 star: 24
- 3 star: 5
- 2 star: 5
- 1 star: 9

⭐⭐⭐⭐⭐ No change from the T4i; not necessarily a bad thing
By D. Alexander TOP 100 REVIEWER on April 23, 2013

This is a Rebel T4i with a better 18-55 kit lens. It’s intended as a drop-in replacement for the T4i, which means it’s the same fast, compact stills camera with a touchscreen that simplifies configuration, image review, and the EOS learning curve. There are better movie cameras. Motion tracking for video, while a vast improvement over DSLRs before the T4i, falls short of many mirrorless bodies.

I’m reviewing it from the perspective of a working professional, which means I’m at least as concerned about what it’s missing as what it has. If you’re new to DSLRs, you’re likely to find this camera an immense upgrade in many ways.

Buy it over mirrorless systems and the T2i/T3i if you want faster shooting and tracking with stills and the immediacy of an optical viewfinder. Choose the SL1 for the most petite size, the 60D for a quicker interface and a deeper buffer for raw files, or the 7D for even better motion-tracking. The T4i alone or with the 18-135 STM is equally compelling if it costs less. Image quality is the same between all the crop bodies. Low-light performance improves with the full-frame 6D and above.

⭐⭐⭐⭐⭐ OMG YOU WON'T REGRET IT <3 <3
By Harley on November 4, 2014

This is the coolest camera ever!! Simply because 1. It has such high quality. 2. This camera is indeed very qualified for beginners to professional photography skills. 3. The box came with more than I expected! 4. You may need to upgrade to a bigger bag/ back pack because everything that came with the camera will not fit in the bag... still okay! Why? Because you won't find a deal this cheap in your lifetime! I love this camera to pieces, I still can't put it down after a week now! You will not regret your purchase and you will LOVE the camera itself!

Source: http://www.amazon.com/Canon-EOS-Rebel-T5-Professional/dp/B00J34YO92/ref=sr_1_2?ie=UTF8&qid=1416336324&sr=8-2&keywords=slr+digital+camera
Motivation

- Lens
- Zoom
- Picture Quality
- UV Filter
- Image Stabilizer

...
Motivation

- I want a summary

Feature-by-feature evaluation

Use all reviews

```
<table>
<thead>
<tr>
<th>Digital_camera_1</th>
</tr>
</thead>
<tbody>
<tr>
<td>picture quality:</td>
</tr>
<tr>
<td>Positive: 253</td>
</tr>
<tr>
<td>Negative: 6</td>
</tr>
<tr>
<td>size:</td>
</tr>
<tr>
<td>Positive: 134</td>
</tr>
<tr>
<td>Negative: 10</td>
</tr>
</tbody>
</table>
```

...
Major Phases

- Finding Features
  - The *lenses* worked great.

- Identifying opinion
  - The lenses *worked* great.

- Producing a summary
  - Extractive, list style
Architecture
Finding Features

- **Frequent Feature Identification** *(Association Mining)*
  - Noun phrases of < 3 words
  - In > 1% of reviews in the corpus
- **Compactness Pruning**
  - Phrase > 3 words apart

“I had searched for a digital camera for 3 months.”
“This is the best digital camera on the market”
“The camera does not have a digital zoom”
Finding Features

- **Frequent Feature Identification** *(Association Mining)*
  - Noun phrases of < 3 words
  - In > 1% of reviews in the corpus

- **Redundancy Pruning**
  - For single words
  - p-support < 3

For example, *manual* by itself may not be an interesting feature but *manual mode* or *manual setting* is a meaningful feature phrase.
Identifying Opinions

- **Opinion Sentences**
  - Include a feature phrase
  - Associate a nearest adjective with features *(Effective opinion)*

  “The **strap** is **horrible** and gets in the way of parts of the camera you need access to.”
Infrequent Features

- Use adjectives to identify infrequent feature

  The pictures are absolutely **amazing**.

  The software that comes with it is **amazing**.
“overall this is a good camera with a really good picture clarity & exceptional close-up shooting capacity”

“the auto and manual along with movie modes are very easy to use, but the software is not intuitive”

In case of a tie use the orientation of the previous sentence
Summary Generation

- **Feature:** picture
  - **Positive:** 12
    - Overall this is a good camera with really good picture clarity
    - The pictures are absolutely amazing – the camera captures the minutest of details
    - After nearly 800 pictures I have found that this camera takes incredible pictures
  
  ... 

  - **Negative:** 2
    - The pictures come out hazy if your hand shake even for a moment
    - Focusing on a display rack about 20 feet away in a brightly lit room during day time, pictures produced by this camera were blurry and in a shade of orange.
Evaluation

- Recall and Precision at each step of feature generation

<table>
<thead>
<tr>
<th>Product name</th>
<th>No. of manual features</th>
<th>Frequent features (association mining)</th>
<th>Compactness pruning</th>
<th>P-support pruning</th>
<th>Infrequent feature identification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Recall</td>
<td>Precision</td>
<td>Recall</td>
<td>Precision</td>
</tr>
<tr>
<td>Digital camera 1</td>
<td>79</td>
<td>0.671</td>
<td>0.552</td>
<td>0.658</td>
<td>0.634</td>
</tr>
<tr>
<td>Digital camera 2</td>
<td>96</td>
<td>0.594</td>
<td>0.594</td>
<td>0.594</td>
<td>0.679</td>
</tr>
<tr>
<td>Cellular phone</td>
<td>67</td>
<td>0.731</td>
<td>0.563</td>
<td>0.716</td>
<td>0.676</td>
</tr>
<tr>
<td>Mp3 player</td>
<td>57</td>
<td>0.652</td>
<td>0.573</td>
<td>0.652</td>
<td>0.683</td>
</tr>
<tr>
<td>DVD player</td>
<td>49</td>
<td>0.754</td>
<td>0.531</td>
<td>0.754</td>
<td>0.634</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>69</strong></td>
<td><strong>0.68</strong></td>
<td><strong>0.56</strong></td>
<td><strong>0.67</strong></td>
<td><strong>0.66</strong></td>
</tr>
</tbody>
</table>
Evaluation

- Opinion Sentence Extraction
  - Precision: 0.693
  - Recall: 0.642

- Sentence Orientation
  - Accuracy: 0.842
Limitation and Future work

- No pronoun resolution
  - “it is quiet but powerful”
- Only adjectives are used as indication of opinion orientation
  - “I highly recommend the camera”
- Strength of opinion not considered
Conclusion

- Mining and summarizing product reviews
- Useful for both shoppers and manufacturers.
- Good accuracy for sentence orientation prediction
Related Study

- **Semantic Classification of Reviews**
  - Supervised learning (Use labeled corpus to classify)

- **Subjective Genre Classification**
  - Topic search (No opinion generation)

- **Text summarization**
  - Mostly used for single document
“the camera is not easy to use”
“it would be nicer not to see little zoom sign on the side which is ugly”