GR-3: Paper Prototyping for Schnap It!

David Hayden, David Way, Joseph Lim, Carl Vondrick
{dshayden, dway, joelim, vondrick}@mit.edu

In order to gain insight into how our user interface would perform on real users, we preformed a variety of paper prototyping experiments on our preliminary interface for Schnap It!. In this document, we present our paper prototyping protocol, report our results, and offer analysis on its impact for building learnable, efficient, and safe user interfaces for visual recognition applications on mobile applications. Our results demonstrate that visual recognition interfaces require both clear textual labels and intuitive annotation modes. We now present these results in detail:

Experiment Overview

We conducted two iterations of paper prototyping to determine how users would react to our mobile phone visual recognition interface. Between iterations, we updated the interface to reflect participant's feedback. For each iteration, we located three subjects who identified themselves as technological innovators. In all cases, number of females = 0 and median age is 23. Each subject volunteered to participate in our experiment and could terminate the experiment at any point. Upon entering our testing room, we presented them with the following briefing:

> **Introduction**
> Looking up products reviews and pricing while at a store is currently tedious: either you thumb-type each product name/serial into Google on your smartphone or you take a picture of each and hope that services like Amazon Remembers will recognize the product. There is no facility for efficient comparison-shopping. We aim to change that with our new mobile application, Schnap It!
>
> **Task and Background Information**
> You're at a bike shop and have set aside several bikes that you're interested in purchasing. Not being a bike expert and not certain that the store is offering a good price, you decide to use Schnap It! to look up and compare online reviews and pricing. You pull out your cellphone to take a picture…

After receiving acknowledgement from the subject that they understood the prompt, we randomly permuted the task list and gave them the first task. We asked the users to perform the following tasks in random order:

1. Find a bike in the store and look up it's reviews.
2. Find a bike in the store and buy it online if it's cheaper.
3. Compare multiple bikes in the store to determine the best one.

After reading the task, we gave them the first screen of the prototype. At task completion, we gave them the next task. We solicited for qualitative feedback as well.
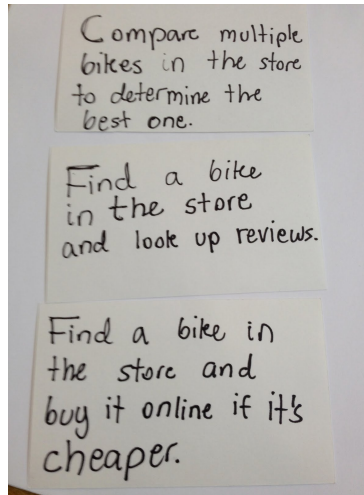
The task cards that we randomly assigned to users.

First Prototype Iteration

We now describe the interface for the first iteration. We sketched this interface on A3 Paper (16.5" x 11.7") using colored markers. Upon opening Schnap It! the user is presented with a live camera view with buttons "Shoot" and "Compare". The Compare Button has an iPhone-style numeric badge, showing how many products can be compared. If it is 0, as is initially the case, then nothing will happen.
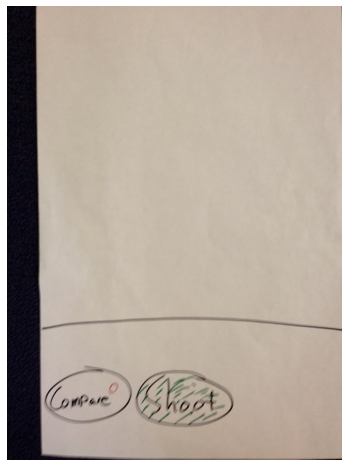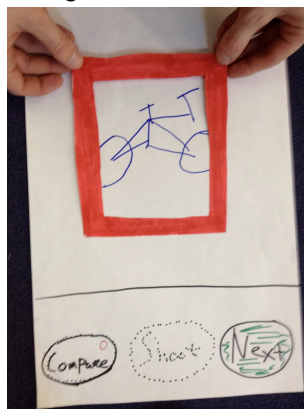

Upon centering one or more products (bicycles in our case), the user is presented with an updated live camera view:
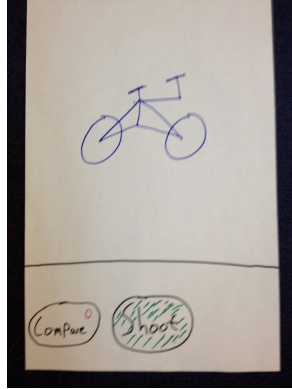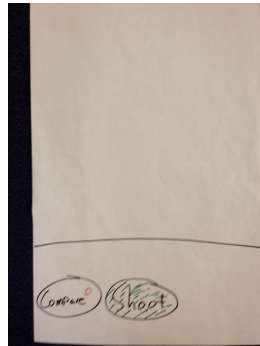
They click "Shoot" and are presented with the "Annotation View". A red, bounding box cut-out is placed at the corner of the screen, ready for them to manipulate. Notice that the "Shoot" button is greyed out.


They are then able to center the bounding box.

Clicking "Next" will take them back to the "Live Camera View", but now the Compare button's badge has a 1 (not shown in this image).

They can repeat the process for one or more bikes. Each time they take a picture of a product, center the bounding box, and click "Next", the Compare button's badge increments. If they attempt to bound multiple objects with a single rectangle, they are presented with an error, "No objects found."
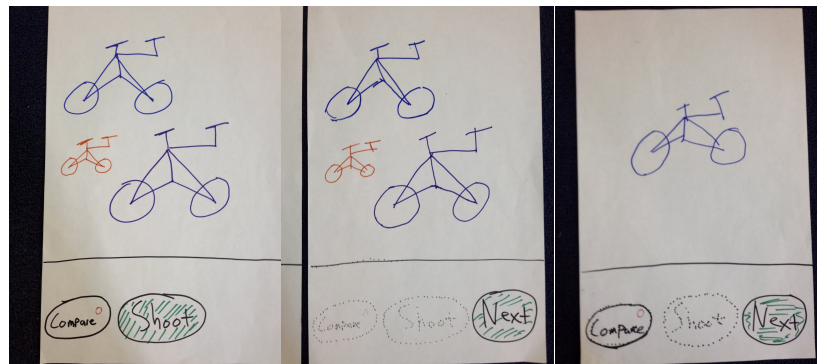
Clicking Compare when it's badge is 1 or greater brings them to the "Product Comparison View" which will have as many products as they successfully imaged and drew bounding boxes around.
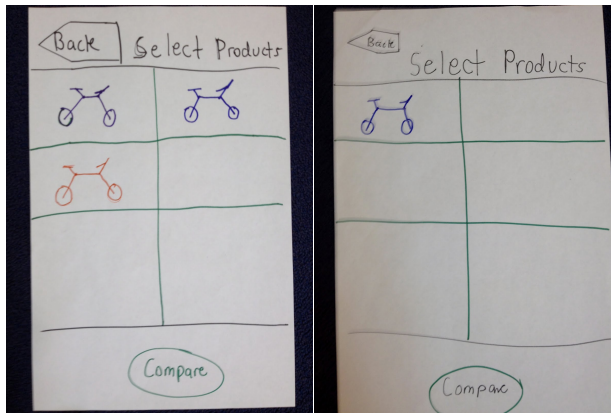
Tapping a product will highlight it; tapping Compare will take them to the "Reviews Shopping Cart View".
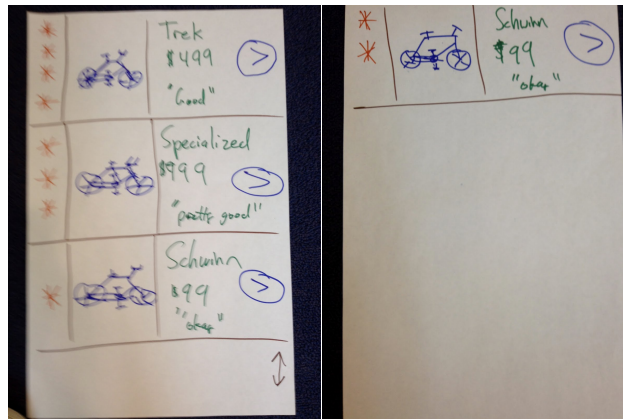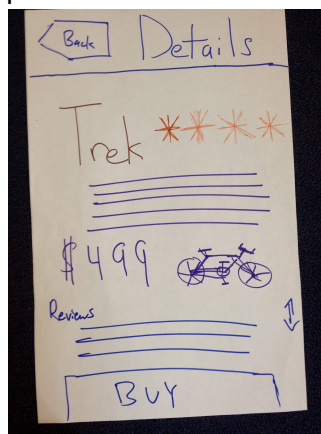

Each cell contains one product with it's name, average rating from reliable review sites, the best price from a reliable online vendor, and a quip from the product's review. From here, users can tap the arrow of any product to see product reviews:

Tapping Back will take them to the "Reviews Shopping Cart View". Tapping "Buy" will open the iPhone's Safari web browser to the vendor that is selling the product for the listed amount.

Observations

Our users provided valuable feedback to the first iteration of paper prototyping. In summary, we found that users were commonly confused by the textual labels of the buttons. Moreover, although some users found the annotation interface intuitive, others were incredibly confused on how to annotate a product.

Tagging by Drawing Enclosing Hull is Confusing: Initially, we provided users with the ability to draw outlines around one or more bike in the same picture with their finger. This was to be compared to the bounding box method. But no users were able to figure out how to annotate in this way. The first utterance from our first user was, "Shit...". They tapped and were presented with small red dots, but never realized that they needed to press and hold to draw. Due to the extreme confusion, we removed this as an annotation method and stuck to the bounding box method. The bounding box method also improves the safety and accuracy of our application. Users can easily modify the bounding box for an accurate description of the object.

Comparison Button is Poorly Labeled: Two of three users were confused when they tapped "Compare" before tapping "Shoot". Since there was nothing to compare, the application did nothing. It was recommended that we grey out Compare when there is nothing to compare. This usability issue could also be solved by some instruction indication.

Live Camera View is Difficult to Model in Paper Prototypes: Almost all users expressed confusion in understanding the difference between the blank camera screen and the camera screen with bikes inside of it. Our paper prototype separated the subtasks pointing the camera at the bike and pressing shoot, where users actually view taking pictures of objects as one whole task. This observed issue would not appear in the implemented prototype, but user confusion became frustration if not immediately explained by the computer. The inefficiency of the design paper prototype and not the actual design of the UI led to user dissatisfaction.

Tap Target Area Must be Enlarged: Two of three users tapped the image or text of a bicycle in the "Reviews Shopping Cart" rather than tapping the arrow. They were frustrated because

the application did nothing. The design of the UI must be more consistent with other platforms. Interaction with table UI's often allow for feedback when users click all fields.

Safari Web Browser Screen Must be Accurate: All users were confused by our somewhat nondescript Safari screen. This was a problem in our prototype's fidelity, rather than an underlying application usability problem. To make the design more efficient, we considered adding a "buy" button earlier in the application. The user would not have to read through reviews to complete the task of buying a bike.

Users Were Confused by Number of Products to Tag: One user was confused by our bounding box implementation; it wasn't clear that they were expected to only center it around one product. They also felt it should be there before they even click "Shoot" to take a picture!

Minimize Number of Steps: Two users forgot their task while in the middle of performing it. To correct this issue in safety, it would help to reduce the number of steps in the application to complete the task.

Extra Buttons Spurs Inefficiency: Users were annoyed that they had to click "Next" between products. All found the interface tedious and rigid. We correct this safety issue by revising button word choice. It would be difficult to improve the tediousness of the application; specifying bounding boxes of objects in general is not the most easy of tasks.

Unnecessary Grid View: One user was confused by the grid view when there was only one product; they thought this was an unnecessary view. We can improve the efficiency of the application in this step.

Revisions

Based on user feedback and our observations of what tended to annoy or confuse users, we made the following changes:

The "Product Comparison View" was eliminated. Any time a product is successfully imaged, it is appended to the "Reviews Shopping Cart". The "Reviews Shopping Cart" now supports removal of products.

The "Live Camera View" has a hole cut out to better simulate the view from a camera. This was an improvement of the paper prototype, and not the design of the UI. In doing so, we were able to observe the problems brought on by the design rather than by the fidelity of the paper prototype.

All views now have a descriptive title that suggests what the user is expected to do. Notably, these are "Schnap It!" for the "Live Camera View", "Tag It!" for the "Annotation View". These titles provide simple, catchy, and exciting instructions. They seemed to tell the user

exactly what to do in the process of the application.

The Compare button is renamed. The Compare button on the "Live Camera View" and "Annotation View" of the previous iteration has been renamed to "Lookup" and is only visible in the "Annotation View" to reduce confusion.

Enlarged target area. The "Reviews Shopping Cart View" will now take users to a product's "Reviews View" if they click anywhere within the cell, rather than if they only click the arrow on the right.

Second Prototype Iteration

After making the changes previously discussed, we built our second prototype. In this version, the user is presented with the "Live Camera View" when they first start the application.
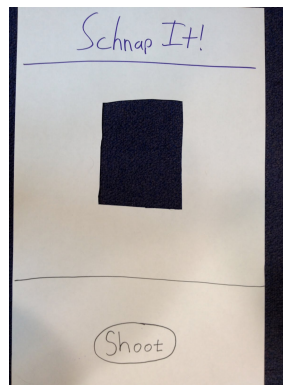

Upon centering one or more bikes, they can click, "Shoot".
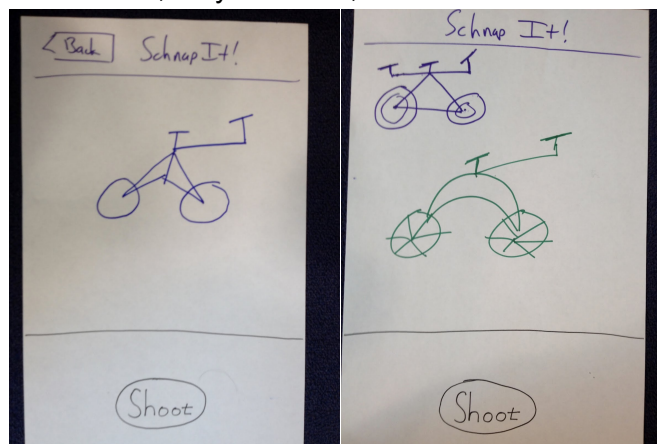
Clicking "Shoot" will take them to the "Annotation View" where they are prompted with "Tag It!" A single red bounding box is placed in the upper corner. There is no misleading 'compare' button.
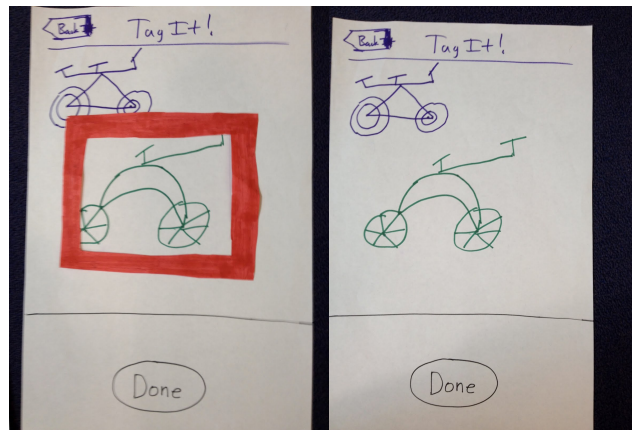
Users can drag the bounding box to cover a single bike. Dragging it around multiple bikes causes an error, "No objects found!" and takes them back to the "Live Camera View". After bounding the object or bike, the user hits done and is prompted to the screen shown below. This screen shows an immovable red bounding box around the objects that have been annotated.
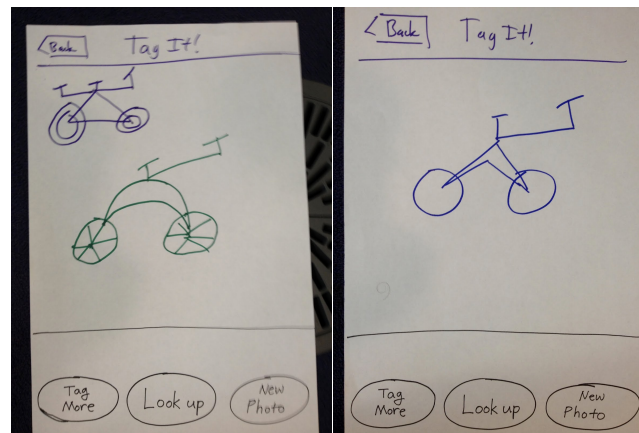

Each time the user clicks "Tag More" they are provided with an additional bounding box, which they can center around other bikes in the image if there are multiple. Note the bounding boxes that have previously been centered on bikes within the same image remain there and cannot be manipulated.

Tapping "New Photo" will take them back to the "Live Camera View", in case they want to take pictures of multiple products which aren't necessarily within the same picture. In this case, a back button appears in the top left of the application on the live camera view screen. By pressing back, the user is presented with the previous picture with corresponding annotated objects screen.

Tapping "Lookup" will take them directly to the "Reviews Shopping Cart View", which will contain

as many bikes as they successfully placed bounding boxes around in one or more images.
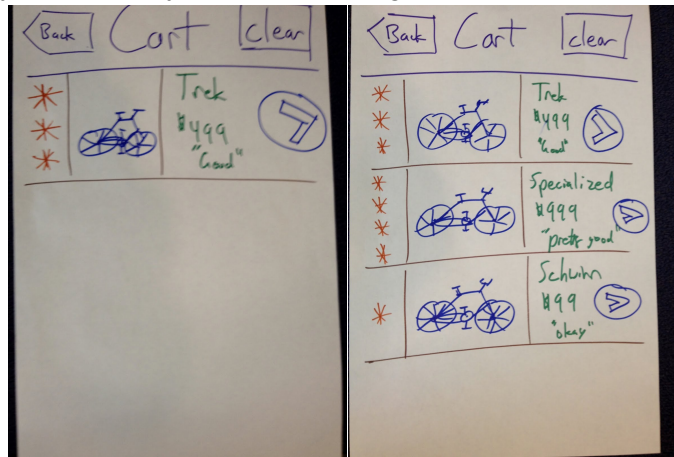

Users are now able to go back to the taking more pictures via the "Back" button. Rather than having to select products to compare through the "Comparison View", they are now more simply able to compare all products at the same time. If something is no longer interesting, they can tap "Clear" and select products to clear.

Tapping any part of a product in this view takes them to the "Product Review View".

And tapping the Buy button takes them to a Safari browser page where they can purchase the product from a merchant's page.
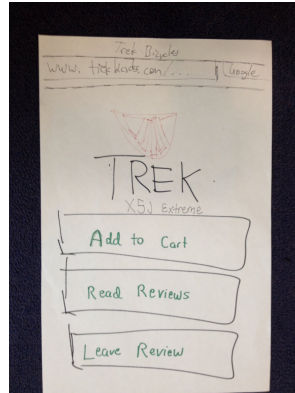
Observations

Overall, the usability of this interface was dramatically improved. Two of three users were able to very quickly navigate through each task, and qualitatively exhibited less frustration than did previous users. The third user ran became confused at several points, in part, we believe, due to language issues. However, there were still some points of failure:

Confusion with Live Camera View: Users found it confusing that we replaced the "Live Camera View" with a static picture of the bikes they centered in the camera, before they tapped "Shoot". This was a problem with prototype fidelity rather than product usability. Midway through the user testing, we stopped using that view; users could then center the bikes in the camera's view and directly tap "Shoot".

Popular Culture Names are not Descriptive: One user of Asian descent was confused by our play on words in the product name, "Schnap It!" He did not realize this was a reference to taking a picture until we explained. The same user did not realize that the bounding box was intended to be placed around a product in the image. Thus, he became confused when he pushed "Lookup" but got an error message, "No product found!" Upon repeating the process, he realized that the bounding box could be moved. Later, when he was asked to compare multiple bikes, he was again confused when the same error message popped up after he dragged the bounding box around multiple products in the image and clicked "Lookup". User claimed that the UI itself was fine, but our choice of props made it confusing (we only had fake bikes, our live camera view was imperfect, the bounding box cutout could not obviously be stretched). This user suggested that we show how many objects have already been labeled.

Worries of Safety by Users: One user was concerned that clicking "New Photo" would throw away all of their existing product images, i.e. it would empty the "Reviews Shopping Cart". The fear of interaction was caused by not clearly describing what the "New Photo" button does. We could improve by choosing more descriptive word choice such as "add new photo". This phrase eliminates the safety issue and adds instruction.

Confusion on Number of Products Identified: Some users seemed to lose track of which and how many objects they had already annotated. Users could have easily checked this by hitting 'look up' and then reviewing which objects have been recognized. However, we could make this process more efficient by keeping a count of how many objects already recognized or a short description such as 'two trek bikes' somewhere. However, the efficiency issue of not keeping track of objects annotated (or tasks) did not seem to be large enough to compensate for.

Future Improvements

In order to address these issues, we plan to make the following changes in our computer prototype:

Clarify Number of Products in Shopping Cart: We should show the number of products the user has already added to the "Reviews Shopping Cart" as a iPhone-style numeric badge on the "Lookup" button of the "Annotation View".

Instructions During Tagging: We need to figure out how to prompt users that they can freehand-draw around an image using one or two words (this method was removed from both iterations of our usability tests because it was caused extreme confusion).

Revise Wording for Foreigners: We should take non-native English speakers into consideration when we use punchy titles/instructions like, "Schnap It!" or, "Tag It!"

Conclusion

We have presented the results of six comprehensive user studies using paper prototypes based on our preliminary user interface for Schnap It!. Our results reveal that designing this interface presents subtle challenges that, if not properly addressed, can severely diminish the learnability, efficiency, and safety of our interface. However, by leveraging low fidelity prototypes, we were able to pinpoint failures before product development and deployment. Since we found these prototypes to be insightful, we plan on conducting more paper prototypes before implementing the computer prototype. Indeed, clear textual labels and intuitive annotation modes are necessary for high performance user interfaces for visual recognition on mobile platforms.