



Empowering blind and low-vision users of
Apple products and related applications

[My account](#) [Submit content](#) [Donate](#) [Help](#) [Log out](#)

Free New OCR App called Voice, feedback appreciated

Submitted by [Shalin Shah](#) on 12 April, 2015 - 04:10 and last modified on 16 April, 2015 - 21:10

App Developer

I am a high school student and app developer and I created an OCR application called Voice that I plan to keep free forever and I would love some feedback so that I can make it as easy to use as possible. I am not blind or low vision, but I really want to help the community. So if you could take the time to give any feedback or suggestions, I would really love it. To use voice, you can simply take a picture of a document or anything with text and Voice will read it for you. Here are some of the features of Voice:

1. Fully compatible with Voice-Over.
2. Field of view report where it detects whether the document is in view, and if it is, it will say "Four corners detected".
3. If the automatic capture feature is turned on in settings, the app will find the document using the camera, and automatically take a picture when the document is in view, without any user interaction.
4. If an image is taken at a slight angle, it will fix the perspective distortion and align the image properly.
5. If the image is of a curved surface, it will adjust the image to straighten it.
6. Book mode which allows the user to take multiple photos, and it will read the photos one after another. The great thing with this feature is that while the first photo is being read, the second is being processed and so on.
7. If the image is too bright, or not bright enough, the app will correct the brightness of the image before processing.

8. Vertical and horizontal column detection for reading different columns in a newspaper; the app does this automatically.
9. Only a 6 megabyte install size.

So those are some of the features of Voice. The photos are only saved temporarily. Once the app has finished reading, the photos are automatically deleted.

I hope you find this application helpful, as I am planning to keep it free forever and just wanted to build it to help people who really need something like this for free. I am open to any suggestions; in fact, please be as brutal as you want so that I can really do some good with this app. Thank you for your time.

Here is the link to the app on the iTunes App Store:

<https://itunes.apple.com/us/app/voice-take-picture-have-it/id903772588?mt=8>

And to contact me personally, please feel free to contact me at shalinvs@gmail.com

Thank you for your time and I hope this would be helpful.

User Options

Hello, how about you use the

Submitted by [burak](#) on 12 April, 2015 - 05:52

Hello, how about you use the apple tts engine so we don't have to have internet for the app to work? Thank you, this is a really good choice.

2 questions

Submitted by [Krister Ekstrom](#) on 12 April, 2015 - 06:28

Any chance of this app going international? I would really love to see this app in Swedish. Also how about adding a video mode? What we, at least some of us need is a way of reading digital displays on things like Microwaves, digital recorders, control surfaces for music etc and it would be great with a realtime video mode for this.

impressive

Submitted by [LaBoheme](#) on 12 April, 2015 - 06:50

pretty impressive. this app doesn't seem to detect page orientation, so if you don't get any result, try reorient the page.

i know you want to keep this app free, "forever" as you put it, but how about making a full blown version and charge for it? by full-blown, i mean being able to save and manipulate the ocr result and maybe adding additional languages? you have at least one buyer here.

Other languages

Submitted by [Jurgen](#) on 12 April, 2015 - 08:11

Hi,

I like the simplicity of this app. But I wasn't able to get it to read. Maybe because I chose a german text? Do you plan to add German or more as an OCR and speak language? I switched on the automatic mode but it seemed not to notice when the edges were visible. Maybe I did something wrong.

All the best for the efforts

Jürgen

Hi this app sounds like it

Submitted by [alex wallis](#) on 12 April, 2015 - 12:46

Hi this app sounds like it has interesting features. However I can't get it to work, I start the app and on the screen first thing I see is a button with no text label. Then I see text that says take a picture and have it read to you, next I see something where you can select 6 pages.

However changing this doesn't do anything going from page 1 to 2 and so on. Tapping the unlabelled button doesn't do anything, I haven't been asked for permission to access my camera or anything. I am using the app on a 6+.

How to get permission to authorized your camera

Submitted by [Marco](#) on 12 April, 2015 - 13:13

Hi Alex

Tapping where is the pages label 1 of 6

Each page include different message

Keep Tapping there until reach page 6 of 6 you will find continue button

After hit this will show some message to allow permission to your camera

Why is internet connection required?

Submitted by [peter](#) on 12 April, 2015 - 13:44

If the OCR is done on the phone, why is an internet connection needed for text to speech?

Once the app determines the text from the OCR VoiceOver should be able to read the result.

Also, one question: Is it possible to save the OCR results in a file that can be exported to other apps or the cloud?

Looks intriguing.

--Pete

The internet connection

Submitted by [Mani](#) on 12 April, 2015 - 14:40

Maybe the internet connection is required because the heavy-duty processing is being done on the server. By Heavy-duty processing I mean the process of converting the image to text. And if that is true then why not do the TTS on the same server? Perhaps that is what the developer is thinking.....

Re: The Internet connection

Submitted by [peter](#) on 12 April, 2015 - 14:47

The author's original post indicated that the OCR process is done on the phone (although it is hard to believe that could be done in the 6 MB size that the author claims). That is why I asked if the internet connection was really necessary for the text to speech part (which presumably

could be handled by VoiceOver).

Also, if the app is free and OCR is being done on a server in the cloud, there must be someone supporting a server, connections, etc. which costs money.

Something a bit strange - Almost too good to be true!

--Pete

New languages coming

Submitted by Shalin Shah on 12 April, 2015 - 14:51

App Developer

Hey Krister,

Yes, in the next or at least one of the upcoming updates, I am planning to add multiple languages (around 40) that are most commonly used around the world.

I agree

Submitted by Mani on 12 April, 2015 - 15:46

The size of the app was what led me to my conclusion. I think the developer owes us an explanation for this.

Explanation for app size

Submitted by Shalin Shah on 12 April, 2015 - 16:21

App Developer

Hey Peter and Mani,

Using auto cropping and computer vision algorithms, I am compressing the image and extracting the essential portion of the image which is the document itself. Everything else in the picture is irrelevant so I am getting rid of it before processing. So the OCR can take place without heavy spacing required. Hence the 6 megabyte size. Also, on the App Store, it even says that the size of the application is 6 MB. Regarding the server, no matter what the cost is to

maintain, I will still keep the app free and try to keep the size under 15 MB. And in the next version, I will make the application work without the need for an Internet connection.

Thank you so much for your suggestions and if you want me to clarify some more things or if you have additional questions, please feel free to contact me.

Re: Explanation for app size

Submitted by [peter](#) on 12 April, 2015 - 16:30

Shalin,

Thanks for your explanations.

It would be nice if you could port the OCR routines over so that they could be run on the phone itself without the need to send the picture to an external server over the net. The newer iPhones should have plenty of horsepower to do that.

Also, as I suggested earlier, it would be nice if the recognized text could be captured in a document which the user could then either save to another app or the cloud to be edited later or read at another time.

Thanks for your efforts. I'm sure this will be appreciated by many people.

--Pete

Thanaks, Shalin.

Submitted by [Mani](#) on 12 April, 2015 - 23:45

Hello Shalin:

Thanks for your explanation. The advantages of delegating the processing on to a server needs to be carefully weighed against local processing. Keep up the good work, Shalin.

mani

Hi Pete,

Submitted by Laszlo on 13 April, 2015 - 10:32

Hi Pete,

Shalin states clearly that the OCR process (the conversion of the image to text) takes place on the phone. That means no image is sent to any server. Right now the resulting text is sent to a text-to-speech server and presumably the synthesised voice is sent back to the phone. Shalin promises to get rid of this in a next update, which would be very very good. I don't really like TTS servers, Voiceover on the phone is much more reliable.

Thanks!

Submitted by Roosterloop Drew on 13 April, 2015 - 13:22

First off, I wanted to say thank you to Shalim for donating his time into a project such as this. We, (The VI community), have not yet seen time donated, into an free OCR app. If you walk away with anything from my post today, let it be graditude and thanks for your time, for the app.

Secondly to the VI community reading this. Let us remember the dev is a high school student, is probably loaded with classes (time is thin), paid for a dev account to even publish the app, and is not blind, and has found us and wants to assist us, again, (for free)

Next up, let us remember the app is in it's inisial release state, and as with any app, there will be bugs to iorn out and changes to be made. I guess I'm just amazed that there hasn't been another post like this one, just saying, "Hey, thanks for your time". All the posts are saying what could be changed.

Regards,

Drew

I haven't tried the app but.

Submitted by brad on 16 April, 2015 - 07:05

I'd just like to say a big thank you for this app. I always find it amazing when an app developer who is sited comes along and tries to help us. So, thank you and good luck with developing this app and school.

Great Version 1.0

Submitted by Jesse Anderson on 16 April, 2015 - 13:35

Hello. I too would like to say well done on the initial release of the app. I've tried recognizing a variety of documents, business cards, etc., with good results. It's a great start.

For anyone having trouble getting good recognition, I found that you need to have the phone in portrait orientation. I tried some wider documents in landscape mode, and got gibberish. Taking the picture in portrait mode worked beautifully.

Here are a couple of suggestions for future updates:

1. Use local TTS via VoiceOver for faster processing.
2. Detect orientation, so the user could take a picture in portrait or landscape modes, or if the user has the paper upside down. A blind user may not know what is the top or bottom of the page.
2. Review and save text to be exported to other apps. If VoiceOver was used for the voice, text could be displayed on the screen, and the user could navigate around the document by touch, the rotor, or other VoiceOver gestures. This would make reading key parts of a page, like a phone number in the middle of a letter, easier.

Great job on the app so far. I could see charging for it.

this sounds amazing!

Submitted by Toonhead on 16 April, 2015 - 16:13

I'm still shocked and amazed that more people haven't said thank you, and are criticizing the app for stuff it can not do. Folks, really. Why must some of you be just so ungrateful? This is a high school student, giving of his time to create an app that could help lots of people. This is an initial release, it's not going to have all the bells and whistles everyone wants at least not yet. Instead of saying what you don't like about the app, how about being positive for once and giving this dude some proper recognition for his time? Some of you just can not be happy with what you have. Ug.

Test results

Submitted by Laszlo on 22 April, 2015 - 08:34

I did a thorough test on Voice. I sent my findings to Shalin. I'm publishing my e-mail here, 'cause I feel like it.

Since then Shalin has reacted to my question in the e-mail by removing the statement about the OCR being done on the phone. Thanks Shalin.

One point: Shalin wants not just "thank you"s, but also feedback about Voice: reporting problems, comments, suggestions etc. Posting such is not criticizing the app, that's important feedback for Shalin.

Hello Shalin,

Man, you've got a great attitude! Caring about accessibility and the needs of the blind, offering your app free, finding the Applevis site and writing about it there and all. That's awesome, congratulations! I'm totally blind and tested your app on my Iphone 5C under IOS 8.1.2. I did the testing without any sighted assistance, relying only on VoiceOver.

Before I share my comments and propositions with you, a question:

On Applevis you state: "The application's OCR is done locally on the phone, ...", while on an information website for the blind called Blindbargains (www.blindbargains.com (<http://www.blindbargains.com>)) in the news section where your app is also announced, the following is stated: "It's built, in part, using the Google Drive API which recently added optical character recognition as part of its available features."

Which statement is correct?

Now my comments:

- I find the first usage "sequence" unintuitive. One has to double-tap on the "page 1 of 6" label to go on, and must keep doing this until "page 6 of 6" appears, and must finally double-tap the "Continue" button. This is hard to guess and quite pointless. A screen with a short informative introduction and a "Continue" button is enough.*
- The main screen is quite okay, but I find having to find the "Done" button after taking a photo unintuitive. I couldn't figure out how "book mode" worked either, I didn't really find clues.*
- The "Filter" label only says "double-tap to change filters". But it doesn't say which filter I am changing to. So currently it's not really accessible.*
- Please consider using VoiceOver as the default TTS engine instead of the Internet-based TTS engine you use now if Voiceover is turned on. I am not an expert on programming for IOS, but I am quite sure there's a dedicated API for sending arbitrary text to VoiceOver for speaking, like the message "four corners detected" in your app. If you do this, you can free your app from requiring an Internet connection (important!), and you can go international easier.*
- Please consider replacing the TTS screen ("Play", "Pause" etc.) with a scrollable read-only text box where you display the recognised text,*

a slider to change pages (if there's more than one) and a "Back" button. This would be a far better solution. There are several gestures in Voiceover to navigate in text boxes by letters, words etc. We could even copy portions of the text out and paste it into other apps.

- If Voice doesn't find any words to recognise in the image, then it should say "no text found" or something similar. Currently it says this when processing multiple images, but not for single images.

- Please consider adding an "auto-orientation" feature to your app, provided that it is entirely your app that does the OCR. Such a feature is essential for the totally blind when the document is rotated 90, 180 or 270 degrees (upside down etc.). Currently Voice doesn't figure out orientation automatically, and that was the reason for most OCR failures during my testing. I fancy figuring out document orientation is not easy at all, so I don't expect this feature soon.

- I find four corner detection quite reliable, and quite helpful.

- If Voice read out anything at all from an image, OCR results were quite breathtaking! E.g. I could identify boxes of medicine with it, and it read the instructions from the box without errors. If it didn't read anything, then it was due to wrong orientation, not to bad photography.

- Unfortunately I didn't find automatic mode reliable at all. When using it Voice produced almost no recognised text, while using manual mode on the exact same document produced good results. In automatic mode After hearing "four corners detected" I had to wait for some random time or move the camera about to hear my phone taking a photo.

- I note that Voice drained my battery surprisingly fast, way more than average. In a test run I used Voice for about half an hour and took 20 or so images with it. My battery level went down from 65% to 36% during this. On the main screen I heard "Flash off", so I presume flashlight was off.

Keep up your good work, motivation and attitude and thanks for your devotion.

All the best to you

Laszlo

regarding comments and constructive criticism

Submitted by **KE7ZUM** on 22 April, 2015 - 14:55

I like the suggestions and constructive criticism. I think the tune head character need to think about that as well. It's not that we are being negative, I use the same technique-in my classes when I teach. You are doing good\, but here's what you need to work on. Works every time.

I have not had a chance to test the app as I'm using the knfb reader but it's always good to have a second tool under your belt and I might try it when visiting the hotel in may if I go on this

trip.

Thanks

Submitted by [Aliya](#) on 23 April, 2015 - 16:49

Hello. At first I would like to say thank you too much, Shalin, for this app. It's very wonderful thing that we get more and more different solutions for OCR.

Unfortunately I didn't have possibility to check out this app because I don't have any text in English at the moment. So it will be very cool if you will add russian in next version of the app.

Thanks one more time for your useful job and good luck!

Has this app been entered

Submitted by [alex wallis](#) on 23 April, 2015 - 17:13

Has this app been entered into the app directory? I have tried searching but haven't found it. The name makes it a little hard to get narrow search results.

Hmm. No I do't think it has

Submitted by [KE7ZUM](#) on 23 April, 2015 - 17:29

Hmm. No I do't think it has been. I guess they are still waiting until the app reaches final stages of testing?

Take care to all.

Voice by Shalin Shah is a

new

Submitted by [Feliciano Godoy](#) on 19 May, 2015 - 11:41

Voice by Shalin Shah is a great app. Features that can be found on paid apps can be found on this free app. Text on screen and sharing options and the ability to copy the text onto the clipboard would make this app golden. Great work can't wait for your future plans!

Regards, Feliciano For tech tips and updates, LIKE www.facebook.com/theblindman12v (<http://www.facebook.com/theblindman12v>) Follow www.twitter.com/theblindman12v (<http://www.twitter.com/theblindman12v>)

[About](#) [RSS Feeds](#) [Newsletter](#) [Links](#) [Developer Resources](#) [Advocacy](#) [Contact](#)

Unless stated otherwise, all content is copyright AppleVis. All rights reserved. © 2015 (<http://www.applevis.com/copyright>) | [Accessibility](http://www.applevis.com/accessibility) (<http://www.applevis.com/accessibility>) | [Privacy](http://www.applevis.com/privacy) (<http://www.applevis.com/privacy>)

Apple and the Apple Logo are trademarks of Apple Inc. © 2015. All Rights Reserved.