

New Technologies for Scanning Books

Posted At : May 6, 2009 1:09 PM | Posted By : Rixon Reed
 Related Categories: Google, scanning, Atiz, Rixon Reed

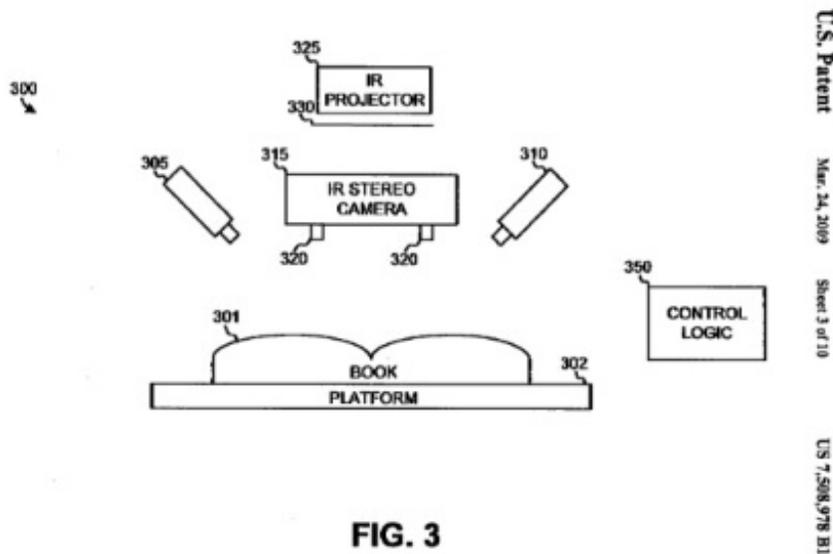


FIG. 3

People love previewing books online. While it's definitely not the same experience as browsing books in a bookstore, this digital equivalent does help give people a sense of a book. This is why we developed our BookTease™ at photo-eye early on, long before Amazon did something similar. While our BookTease gives a good picture of the layout of a book, it's not meant to capture the complete content, nor does it put the text into a searchable database.

My friend Laurie Tümer turned me on to a recent NPR blog post about the patent granted to Google for their new scanning book technology. The Google Book Search project's extraordinary goal is to digitize nearly every book ever published, allowing a person to search for content easily and thoroughly. Think of it as the Google Search engine for books. Apart from the significant copyright issues, there are other serious, technological problems to overcome. Google needed to develop an incredibly efficient system to scan each page of millions of volumes.

Scanning book technology has been around for a number of years with [Atiz](#) being one of the pioneers. Atiz uses a V-shaped apparatus to prevent damage to the books. Other companies apparently destroy the books by cutting each page out sheet by sheet.



Google's technology uses an infrared camera to detect the three-dimensional shape of an opened book. Their proprietary software corrects for the distortion of the type, giving an accurate picture of the information on each page.

Read more here: [NPR blog](#)

But perhaps there's an even better way of doing this. Take a look at this interesting comment by Daniel Mitchell on the NPR blog.

"Do you need to open the book at all? Scan it like a MRI - read the book from the outside with a broad spectrum of different wavelengths and focal points. Create a full 3-D image of the "book" - paper, ink and all. Once finished scanning - hand the data over to a computer algorithm to process the information to recreate the text, pages, textures and all the subtleness of the books content. No damage to the book - no need to turn pages."

Makes sense, though MRI technology may not yet be able to resolve the variability of paper and ink. But that would be a fantastic use of technology!