

Smart File Cabinet Prototype

Teller, Karger, Glass, Dec 2002

Basic idea is to build a device/system that takes paper documents as input. Each document fed to the device is scanned, OCR'ed, and classified as one of (say) a hundred common document types: letter, bill, bank statement, receipt, coupon, cartoon, advertisement, etc.

The system stores a scanned image of the document, along with all OCR keywords extracted. Additionally, any utterance the user makes at the time of input is parsed and attached to the document as metadata.

A video camera watching the sheet feeder acquires a low-res image of each document as it is handed to the device, and a microphone captures user utterances. Thus, in rapid succession (without waiting for the scanning process, which can be slow), the system can support a scenario like this:

[User places document on sheet feeder.]
User utters, "This is the gas bill."

[User places document on sheet feeder.]
User utters, "This is the receipt for Mary's gift."

[User places document on sheet feeder.]
User utters, "This is that haircut coupon."

[User places document on sheet feeder.]
User utters, "This is that memo from MIT benefits."

Later the user can use Haystack to retrieve anything that she filed.

