

FACMOVEL: Aplicação para Digitalização, Detecção e Classificação Automática de Recibos em Dispositivos Móveis

Miguel Batista¹, Pedro Morais¹, Rui Jesus² e Pedro Jorge²

¹Tekever

²M2A, Instituto Superior de Engenharia de Lisboa
Rua Conselheiro Emídio Navarro, n.º1, 1959-007 Lisboa
greavier@gmail.com, morais.pedro@gmail.com,
rjesus@deetc.isel.ipl.pt, pjorge@deetc.isel.pt

Abstract. Ao longo dos últimos anos tem-se assistido a uma rápida evolução dos dispositivos móveis. Essa evolução tem-se vindo a sentir no poder de processamento, na resolução e qualidade das câmaras e na largura de banda das redes móveis de nova geração. Assim, com cada vez mais “computadores de bolso” surge a oportunidade de criar aplicações que conseguem converter informação do mundo físico em informação digital que depois pode ser processada e utilizada de forma eficiente (e.g., um leitor de códigos QR ou um OCR). Este artigo descreve uma aplicação para digitalização, classificação e gestão automática de recibos em dispositivos móveis. A aplicação utiliza técnicas de processamento de imagem, classificação e realidade aumentada para realizar esta tarefa. São também apresentados resultados de testes efectuados para validação da aplicação.

Keywords: Aplicações em Dispositivos Móveis, Processamento de imagem, Reconhecimento óptico de caracteres, Realidade Aumentada.

1 Introdução

Actualmente, alguns processos ainda dependem fortemente da utilização do suporte em papel apesar de já existir, em alguns casos, uma transição inicial para o suporte digital. Por exemplo, a geração de recibos é um processo muito ineficiente pois este é criado a partir de informação digital que depois o utilizador necessita de converter novamente para formato digital para fazer um relatório de despesa, fazer o IRS, etc. Além de ser incómodo ter de reconverter a informação do recibo para digital, estes ainda têm a desvantagem de se poderem perder ou deteriorar com facilidade.

Segundo um estudo recente de mercado [1], nos Estados Unidos da América o número de utilizadores de *smartphones* já ultrapassou o número de utilizadores dos chamados *feature phones*. Estes dispositivos móveis avançados permitem o desenvolvimento de soluções que utilizam tecnologias como serviços Web, *Near Field Communication* (NFC) ou *Bluetooth* para transferir informação do recibo entre os vários recibos (fonte e receptor), mas apesar de elegantes e exequíveis, estas

soluções envolvem a actualização dos terminais de venda, o que poderá torná-las dispendiosas. A solução apresentada neste trabalho procura aproveitar os recursos dos dispositivos móveis para converter o actual recibo em papel para formato digital. Esta abordagem, apesar de manter a impressão do recibo em papel, mantém a infraestrutura actual.

Actualmente existem algumas soluções no mercado para a extracção de informação e classificação de recibos, mas nenhuma dá ao utilizador uma forma automática e exequível em tempo útil, de extracção de informação e classificação dos recibos. Como exemplos, a aplicação *Expensify* [2] depende inteiramente de processamento remoto com custos associados e com um tempo de processamento de até 10 minutos, ou um dia útil caso o processo automático falhe, a aplicação *Shoboxed Receipt Tracker* [3] também depende de processamento remoto e devolve os resultados em até um dia útil, e a aplicação *ProOnGo Receipt Reader* [4] não faz qualquer tipo de extracção de informação ou classificação, apenas faz um processamento de imagem para possibilitar o armazenamento do recibo em formato digital.

2 FACMOVEL

A aplicação descrita neste artigo é composta por 4 blocos: 1 - Aquisição e Detecção, 2 - Extracção de Informação, 3 - Classificação e 4 - Interface Gráfica.

2.1 Aquisição e Detecção de Recibos

Este processo, detalhado na Figura 1 começa por aplicar algumas operações de processamento de imagem de forma a encontrar os contornos do recibo. De seguida, é realizada uma transformação homográfica para corrigir qualquer efeito de perspectiva e rotação. No final a imagem é cortada na área do recibo para obter a imagem final.

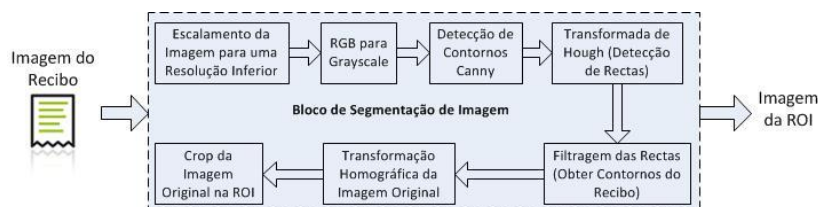


Figura 1 - Diagrama de blocos da Aquisição e Detecção do Recibo

2.2 Extracção de Informação

Este processo centra-se na extracção e tratamento do texto do recibo para posterior processamento. A Figura 2 resume os blocos constituintes deste passo.

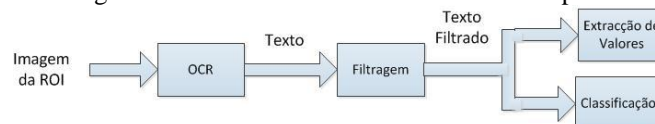


Figura 2 – Diagrama de blocos da Extracção de Informação

2.2.1 Reconhecimento Óptico de Caracteres

O reconhecimento automático de caracteres é realizado com base no OCR *Tesseract* [5]. Este processo é realizado com recurso a um servidor, pois o OCR no dispositivo móvel tem um tempo de execução muito longo, que dependendo das condições da imagem, pode variar entre aproximadamente 15 segundos a mais de 2 minutos. Este processo no servidor demora cerca de um segundo, sendo que o tempo restante é o tempo de transferência, e este depende da dimensão da imagem final e da qualidade da ligação.

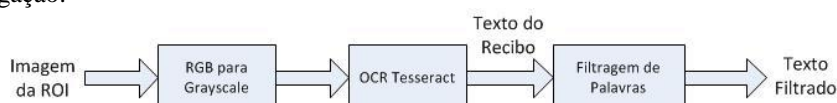


Figura 3 – Diagrama de blocos do Reconhecimento Óptico de Caracteres

2.2.2 Extracção de Valores

A extracção de valores do recibo recorre a colecções de palavras (*bag of words*) para procurar cada valor. Por exemplo, para detectar o valor do Número de Identificação Fiscal são pesquisadas palavras como “NIF”, “NIPC”, ou “N Cont.”. A distância de Hamming também é utilizada em conjunto com esta técnica, pois o OCR pode ter erros e os termos pesquisados podem não estar exactamente na forma que se espera.

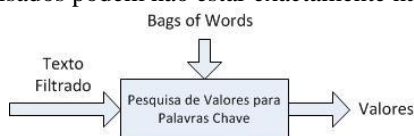


Figura 4 – Diagrama de blocos da Extracção de Valores

2.3 Classificação do Recibo

A classificação é realizada com base no algoritmo de Naive Bayes [6]. É utilizada uma aprendizagem incremental, uma vez que o conjunto de treino (inicialmente vazio) vai aumentando à medida que os recibos vão sendo processados. Após a classificação automática, o utilizador poderá validar, alterar ou adicionar a classe atribuída ao recibo, através da interface gráfica. No final deste processo, a informação do novo recibo é adicionado ao conjunto de treino e o processo de estimação das distribuições é repetido com base nesta actualização (ver figura 5).

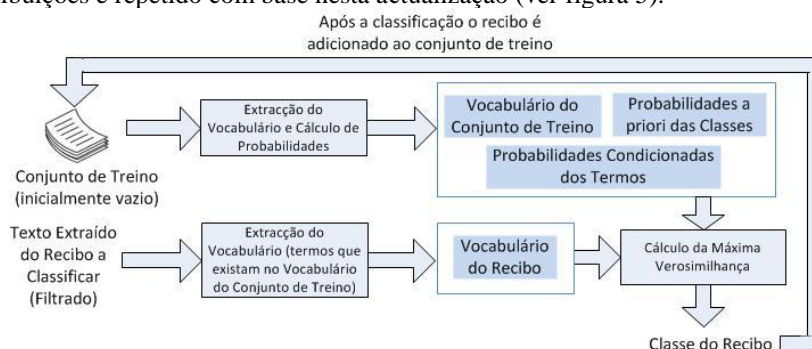


Figura 6 – Diagrama de blocos da Classificação do Recibo

3 Resultados Experimentais

A aplicação foi desenvolvida para um *smartphone* da Apple, modelo iPhone 4, sistema operativo iOS 5, processador 1 GHz Cortex-A8 e câmara frontal de resolução 5 MP (2592 x 1936 pixéis). A sequência de processos necessários para apresentar a componente de realidade aumentada em tempo real durante a detecção do recibo tem uma duração total de 340 ms, sendo a transformada de Hough e filtragem de rectas a operação que demora mais (264 ms). Assim, com um tempo total de processamento de cerca de 340 ms conseguem-se processar aproximadamente 3 imagens por segundo. Para a componente de processamento do recibo, são necessários os 340 ms do processo anterior, o tempo da transformação homográfica que em média durou 9 segundos, o tempo da transferência da imagem para o servidor, que depende da dimensão da imagem e qualidade de ligação, e tempo de OCR no servidor. A duração da extracção de valores e classificação do recibo é inferior a um segundo.

4 Conclusões e Trabalho Futuro

A implementação actual mostra que é possível desenvolver o sistema proposto num dispositivo móvel. Todo o processo desde a aquisição até à classificação é feito em tempo útil, sendo o OCR o processo mais moroso. Testes preliminares mostram que o recurso a um servidor para efectuar este passo, diminui consideravelmente o tempo de processamento, sendo importante para que a aplicação tenha uma melhor usabilidade. Como trabalho futuro, para melhorar a aplicação desenvolvida, dever-se-á investir no aperfeiçoamento da interface com realidade aumentada utilizando por exemplo elementos 3D. Para tornar a aplicação mais fácil e intuitiva de utilizar deverão ser realizados estudos de usabilidade com vários utilizadores, de modo a apresentar uma melhor interface. Finalmente, para tornar a aplicação mais útil, dando mais informação ao utilizador, poder-se-á implementar um sistema de gestão e organização de modo a conseguir organizar os recibos por classe, fazer somas parciais por data, categoria ou entidade emissora do recibo, somas totais, entre outras possibilidades.

Referências

- [1] America's New Mobile Majority: a Look at Smartphone Owners in the U.S. [Online]. <http://blog.nielsen.com/nielsenwire/?p=31688>
- [2] Expensify app. [Online]. <https://www.expensify.com/mobile>
- [3] Shoeboxed Receipt Tracker. [Online]. <https://landing.shoeboxed.com/iphone/1/>
- [4] ProOnGo Receipt Reader. [Online]. <http://www.proongo.com/b/receipt-reading.php>
- [5] Tesseract OCR. [Online]. <http://code.google.com/p/tesseract-ocr/>
- [6] Naive Bayes. [Online]. <http://cis.poly.edu/~mleung/FRE7851/f07/naiveBayesianClassifier.pdf>