

# Correção da Orientação de Imagens de Páginas de Documentos

Gabriel Sichelero<sup>1</sup>

<sup>1</sup>Universidade Tecnológica Federal do Paraná UTFPR - Departamento Acadêmico de Informática  
Av. Sete de Setembro 3165, Rebouças, Curitiba - PR

`gabriel.sichelero@hotmail.com`

## 1. Introdução

Atualmente existe uma crescente necessidade de digitalizar documentos físicos para facilitar o armazenamento e a distribuição de informações. Este processo de digitalização normalmente envolve a captura de uma fotografia ou escaneamento do documento, resultando em uma imagem digital, a partir dessa imagem são extraídas as informações contidas no documento. No entanto, para melhorar a eficiência da extração dessas informações, é preferível que as páginas contidas nas imagens estejam alinhadas horizontalmente e verticalmente, na orientação que o documento foi escrito originalmente.

Mais especificamente o problema de alinhamento das páginas de um documento consiste em encontrar a ângulação da imagem digitalizada em relação à orientação original do documento e aplicar uma rotação para corrigir essa ângulação.

Este estudo analisa os métodos existentes na bibliografia para abordar este problema. Além disso, será definido quais métodos são mais adequados para serem futuramente analisados em profundidade, adaptados e implementados para a resolução deste problema.

## 2. Revisão Bibliográfica

O alinhamento de imagens é um problema que vai além do escopo de documentos digitalizados, existem métodos para alinhamento de imagens em geral, como o RANSAC (Random Sample Consensus) [Shen et al. 2020], que faz o alinhamento de duas imagens através da identificação de características em comum entre elas e a aplicação de uma transformação geométrica para alinhá-las. A desvantagem deste método para o alinhamento das imagens de documentos é que nem sempre teremos a imagem do documento com a orientação correta para fazer a comparação.

Para o problema mais específico da correção da orientação de imagens de documentos existem abordagens mais tradicionais que utilizam métodos de processamento de imagens para extrair características das imagens, como a Transformada de Hough [Hinds et al. 1990] ou a Transformada de Fourier [Fabrizio 2014] para a detecção de linhas paralelas e com isso determinar a orientação da imagem.

Além disso, existem métodos que utilizam redes neurais convolucionais treinadas com imagens de documentos com diversas orientações para tentar prever a diferença entre a orientação da imagem e a orientação correta do documento [Akhter and Rege 2020]

[Aggarwal et al. 2021]. Esses métodos são treinados usando bases de dados que contêm imagens de documentos com a orientação distorcida, como o ICDAR 2013 Document Image Skew Estimation Contest (DISEC 2013) [Papandreou et al. 2013] e não necessitam de uma imagem de referência do documento com a orientação certa para fazer a correção.

### 3. Método

Para definir quais métodos irão ser analisados com mais profundidade e desenvolvidos vamos considerar os seguintes critérios: primeiramente, vamos considerar que não teremos a imagem do documento com a orientação correta para usar de referência e fazer a comparação, portanto métodos que necessitam de uma imagem de referência não serão considerados. Além disso, como existem bases de dados disponíveis com imagens de documentos com a orientação distorcida, podemos considerar métodos baseados em aprendizagem profunda que precisam ser treinadas com essas bases de dados.

Com base nesses critérios, podemos implementar um método recente que utiliza redes neurais convolucionais [Akhter and Rege 2020] e comparar o desempenho com uma implementação atual de um método tradicional baseado na Transformada de Hough [Ahmad et al. 2021].

Por fim, para validar a eficácia dos modelos, serão realizados testes com imagens de documentos com a orientação distorcida de bases de dados públicas, como o ICDAR 2013 Document Image Skew Estimation Contest (DISEC 2013) [Papandreou et al. 2013].

### Referências

- Aggarwal, S., Gaur, S. S., and Manju (2021). Text document orientation detection using convolutional neural networks. In Sharma, H., Saraswat, M., Kumar, S., and Bansal, J. C., editors, *Intelligent Learning for Computer Vision*, pages 153–164, Singapore. Springer Singapore.
- Ahmad, R., Naz, S., and Razzak, I. (2021). Efficient skew detection and correction in scanned document images through clustering of probabilistic hough transforms. *Pattern Recognition Letters*, 152:93–99.
- Akhter, S. S. M. N. and Rege, P. P. (2020). Improving skew detection and correction in different document images using a deep learning approach. In *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, pages 1–6.
- Fabrizio, J. (2014). A precise skew estimation algorithm for document images using knn clustering and fourier transform. In *2014 IEEE International Conference on Image Processing (ICIP)*, pages 2585–2588.
- Hinds, S., Fisher, J., and D’Amato, D. (1990). A document skew detection method using run-length encoding and the hough transform. In *[1990] Proceedings. 10th International Conference on Pattern Recognition*, volume i, pages 464–468 vol.1.
- Papandreou, A., Gatos, B., Louloudis, G., and Stamatopoulos, N. (2013). Icdar 2013 document image skew estimation contest (disec 2013). In *2013 12th International Conference on Document Analysis and Recognition*, pages 1444–1448.

Shen, X., Darmon, F., Efros, A. A., and Aubry, M. (2020). Ransac-flow: Generic two-stage image alignment. In Vedaldi, A., Bischof, H., Brox, T., and Frahm, J.-M., editors, *Computer Vision – ECCV 2020*, pages 618–637, Cham. Springer International Publishing.