



Universidade Federal de Pernambuco  
Centro de Informática

Graduação em Ciência da Computação

# **Reconhecimento de Imagens de Documentos Através de Características Invariantes a Transformações**

João Gabriel Santiago Mauricio de Abreu

Proposta de Trabalho de Graduação

**Orientador:** Veronica Teichrieb

**Coorientador:** João Marcelo Teixeira

Recife  
Abril de 2017

# Sumário

<b>Contexto</b>	<b>3</b>
<b>Objetivos</b>	<b>5</b>
<b>Cronograma</b>	<b>6</b>
<b>Referências</b>	<b>7</b>
<b>Assinaturas</b>	<b>8</b>

## Contexto

Na última década, as câmeras digitais se tornaram dispositivos amplamente acessíveis, devido a estarem comumente presentes em smartphones, que se tornaram essencialmente ubíquos nesse período. Além dessa grande difusão, a qualidade das câmeras presentes nesses aparelhos também melhorou consideravelmente, de modo que, atualmente, a maioria das pessoas carrega sempre consigo uma câmera 720p.

Essa grande presença de câmeras digitais entre as pessoas possibilita e incentiva a existência de aplicações e algoritmos que se baseiam em imagens capturadas pelos usuários, especialmente aplicações direcionadas para dispositivos móveis. Uma necessidade que foi gerada por essas novas possibilidades é o reconhecimento de documentos ou páginas de texto a partir de fotos capturadas por uma câmera. Uma grande motivação para esse tipo de reconhecimento é a recuperação de conteúdo adicional relevante ao documento, como figuras ou anotações. Além disso, documentos também podem ser utilizados como identificadores, funcionando essencialmente como códigos de barras ou QR. Finalmente, também é possível renderizar sobre o documento algum conteúdo virtual relevante, respeitando sua localização espacial.

O problema descrito é chamado busca por imagens de documentos (*document image retrieval*) e consiste em, dada uma imagem de um documento, encontrar as imagens dos documentos mais similares dentre as presentes em um banco de dados (Fig 1). Este trabalho aborda uma versão mais específica deste problema, chamada busca por imagens de documentos baseadas em câmera (*camera-based document image retrieval*), no qual as imagens fornecidas pelos usuários foram capturadas por uma câmera. Esta especificação adicional, de que as imagens foram fotografadas, implica em uma série de problemas adicionais que devem ser resolvidos, principalmente: baixa resolução, ajuste do foco da câmera, *motion blur*, iluminação heterogênea e, sobretudo, distorção projetiva.

Atualmente, existem alguns algoritmos eficientes e bem difundidos que abordam esse problema, apresentando estruturas semelhantes entre eles [1][2][3]. Entretanto, esses algoritmos possuem restrições consideráveis de consumo de memória que limitam seu uso em várias plataformas a somente bancos de dados pequenos.



Fig. 1: Exemplo de par de imagens de um mesmo documento. Na esquerda, a imagem contida no banco de dados. Na direita, a imagem capturada pelo usuário.

## Objetivos

O objetivo deste trabalho é realizar uma pesquisa acerca do estado da arte no problema de busca por imagens de documentos baseadas em câmera. Então, a partir desta pesquisa, implementar uma solução adaptada que seja viável para dispositivos móveis. Não é parte do escopo deste projeto implementar aplicações para esses dispositivos, porém levaremos em conta as restrições de memória dessas plataformas para que a solução funcione como uma biblioteca que possa ser utilizada nelas. Além disso, pretende-se também que essa biblioteca forneça informação suficiente para que seja possível *aumentar* o documento reconhecido através de Realidade Aumentada.

A solução implementada será avaliada de acordo com alguns indicadores de desempenho, sendo eles: acurácia do reconhecimento de documentos, quantidade de frames por segundo e consumo de memória.

# Cronograma

Este trabalho de graduação será executado seguindo o seguinte cronograma.

Atividades	Março	Abril	Maiο	Junho	Julho
Revisão bibliográfica	■	■	■		
Elaboração da proposta		■			
Implementação e avaliação		■	■	■	
Escrita da monografia			■	■	■
Preparação da apresentação					■
Apresentação final					■

## Referências

- [1] T. Nakai, K. Kise, and M. Iwamura. “Use of affine invariants in locally likely arrangement hashing for camera-based document image retrieval”. Lecture Notes in Computer Science (7th International Workshop DAS2006), vol. 3872, pp. 541–552, 2006.
- [2] T. Nakai, K. Kise, and M. Iwamura. “Camera based document image retrieval with more time and memory efficient LLAH”. Second International Workshop on Camera-Based Document Analysis and Recognition, pp. 21-28, 2007.
- [3] Q. B. Dang, V. P. Le, M. M. Luqman, M. Coustaty, D. C. Tran, et al. “Camera-based document image retrieval system using local features - comparing SRIF with LLAH, SIFT, SURF and ORB”. International Conference on Document Analysis and Recognition, pp. 1211-1215, 2015.

# Assinaturas

---

Veronica Teichrieb  
Orientador

---

João Gabriel Santiago Mauricio de Abreu  
Aluno