



**FACULDADE DE ENGENHARIA  
ELÉTRICA E DE COMPUTAÇÃO**

# **EE017 - Estágio em Empresa**

## **Relatório Final de Estágio**

**Aluno:** Carlos André Nunes  
**E-mail:** [carlosandren96@gmail.com](mailto:carlosandren96@gmail.com)

**RA:** 154964

**Supervisora:** Vanessa Testoni  
**E-mail:** [vanessa.t@samsung.com](mailto:vanessa.t@samsung.com)  
**Telefone:** (19) 99703-8842

**Empresa:** Samsung Eletrônica da Amazônia - SRBR  
**Telefone:** (19) 3348-0450

---

Assinatura do Aluno

---

Assinatura da Supervisora

Campinas - 2018

## Introdução

O programa de estágio na SRBR (Samsung R&D Institute Brazil) teve início no dia 11 de agosto de 2018 e está previsto para encerrar, pelo contrato vigente, no dia 10 de Março de 2019. O aluno atua no setor de P&D da empresa, tendo sido alocado em dois times diferentes durante o semestre. Na primeira metade do semestre, na equipe de Media Standards - Advanced Technology, sob supervisão da Dra. Vanessa Testoni.

Durante a segunda metade do semestre, na equipe de Artificial Intelligence - Advanced Technology, sob supervisão do Dr. Otávio Augusto Bizetto Penatti.

Sua carga horária é de 30 horas semanais, cumprindo 6 horas de estágio por dia, seja dentro da própria empresa, seja no Laboratório de Pesquisa Colaborativa localizado no Parque Científico e Tecnológico da Unicamp.

## Contexto do trabalho

Ao se acompanhar a tendência na evolução dos formatos de mídia atualmente, é claramente perceptível o interesse das companhias do ramo nas áreas de realidade aumentada (do inglês, *augmented reality* - AR) e realidade virtual (do inglês, *virtual reality* - VR), que exigem um aperfeiçoamento nas técnicas de amostragem, processamento e renderização de imagens 3D. Percebendo essa tendência, o grupo JPEG<sup>1</sup> (*Joint Photographic Experts Group*), uma cooperativa de trabalho associada à Organização Internacional de Normalização (do inglês, ISO) responsável pelo desenvolvimento de codificadores de imagem, criou em 2015 um grupo separado de trabalho denominado JPEG-Pleno com a finalidade de prover uma plataforma padronizada para a representação de novas modalidades de imagens, tais como *light-fields* e *point clouds*.

*Light-Fields* é um formato de mídia que tem sido muito explorado recentemente pela sua capacidade de gerar com alta qualidade imagens 3D e permitir edições em imagens que antes não eram hábeis, como alteração do foco.

Essa capacidade do *light-field* vem do fato de ela explorar uma versão simplificada da função plenóptica, explorando não apenas intensidade ou direção dos raios de luz mas também o ângulo. Isso permite com que, caso se tenha um dispositivo capacitado, um indivíduo obtenha através de uma simples foto, todo o contexto espacial de um local em três dimensões. Entretanto, a obtenção de mais informações por foto acaba por gerar um entrave no desenvolvimento dessa mídia, que é a quantidade de dados gerada.

---

<sup>1</sup> <https://jpeg.org/jpegpleno/index.html>



**Figura 1. Câmera raytrix R11 com capacidade de capturar *light-fields*.**

Devido ao fato de apresentar um volume de dados bastante superior em relação às câmeras convencionais - os arquivos gerados por elas ocupam um espaço considerável em memória, chegando à 50MB em uma única foto tirada e ultrapassando 400MB após passar por tratamento -, diversos grupos discutem nas reuniões do JPEG-Pleno maneiras de reduzir a quantidade de informação explorando certas redundâncias e, assim, criar um protocolo de compressão de dados que seja o mais eficiente possível, com baixa perda na qualidade da imagem final.

Uma outra tendência cujo crescimento é visível, é das empresas de buscar cada vez mais gerar conteúdo personalizado e direcionado para os consumidores. Isso se torna possível, por exemplo, na análise das páginas na internet que determinado indivíduo costuma acessar ou que tipo de entretenimento o consumidor busca.

Uma forma muito simples porém muito efetiva através da qual as empresas têm buscado realizar isso é pelo desenvolvimento de códigos para reconhecimento de textos e imagens nos conteúdos acessados pelos consumidores. Isso tem se tornado cada vez mais fácil graças à constante otimização e popularização pela qual passam os algoritmos que se baseiam em visão computacional. O uso de algoritmos de reconhecimento permitem às empresas entenderem que tipo de conteúdo os usuários estão buscando e dessa forma oferecer um *marketing* mais personalizado a eles.

Apesar de haver um certo receio por parte da população em ter sua privacidade “violada” - uma vez que todo conteúdo acessado pode estar sendo constantemente avaliado por um algoritmo computacional -, a possibilidade de se ter um conteúdo mais personalizado para aqueles que estão constantemente na internet acaba sendo uma opção extremamente atrativa e benéfica tanto para as empresas quanto para os consumidores.

## **Objetivos do estágio**

A proposta do programa de estágio inicialmente foi inserir o aluno no contexto de *media standards*, especialmente voltado à codificação de vídeos ligados a área de realidade virtual (VR), realidade aumentada (AR) e realidade mista (MR) ao trabalhar com *light-fields*.

Ao longo do semestre o aluno participou de dois projetos. Primeiramente, trabalhou junto de uma equipe da Universidade Federal do Rio de Janeiro na construção de um codificador de *Light-Fields*. Neste projeto, o aluno realizou testes de códigos para compressão de imagens e analisou os resultados obtidos para averiguar a qualidade do

algoritmo desenvolvido, em virtude de se estabelecer futuramente uma possível contribuição para a criação do próximo formato para compressão de imagens.

Posteriormente, o aluno foi alocado a outro projeto, com o intuito de ajudar no desenvolvimento de novas ferramentas para o Sistema Operacional Tizen<sup>2</sup>, um sistema operacional com *kernel* Linux, que teve grande contribuição da Samsung no seu desenvolvimento. Neste projeto, o aluno tem como propósito auxiliar no desenvolvimento de códigos para o reconhecimento de texto em imagens e vídeos.

Assim, o estágio teve como objetivos:

- Implementar algoritmos de codificação;
- Desenvolver ferramentas para apoiar atividades de pesquisa;
- Auxiliar em provas de conceito para tecnologias propostas.

## Metodologia

Inicialmente, para introduzir o aluno ao campo de trabalho, foram feitas intensas pesquisas sobre o estado da arte das tecnologias e *standards* atuais, além da leitura de artigos sobre processos ligados à compressão de mídias digitais e *Light Fields*.

### **1) Testes de códigos para compressão de imagem**

Com o propósito de contribuir com os trabalhos atualmente realizados no grupo JPEG Pleno, a Samsung firmou uma parceria com pesquisadores da UFRJ para o desenvolvimento de um código para compressão de imagens. Caso aceito, o código passaria a integrar o próximo padrão internacional para compressão de imagens.

O aluno ficou encarregado de executar testes com o código desenvolvido para a compressão e descompressão de conjuntos de *datasets* fornecidos pelo grupo de trabalho do JPEG Pleno. Serão então calculadas métricas de qualidade para averiguar a qualidade do algoritmo desenvolvido. Esses resultados virão a ser reportados futuramente durante reuniões do grupo de trabalho para fins de comparação com o que está sendo feito por outros membros do comitê.

O código para o cálculo das métricas de qualidade foi desenvolvido em Matlab e o código para compressão dos *datasets* foi desenvolvido em *shell script*.

### **2) Desenvolvimento de códigos para reconhecimento de texto**

Durante o semestre, o aluno foi realocado para outro time dentro da Samsung onde ele ajudou no desenvolvimento de códigos para reconhecimento de textos em imagens e vídeos que posteriormente serão integrados ao ambiente Tizen, desenvolvido pela empresa para seus dispositivos eletrônicos.

Foi dado ao aluno a tarefa de implementar o reconhecimento de texto em vídeos. Para isso, foram feitos testes e estudos para se determinar o fluxo de execução do código

---

<sup>2</sup> <https://www.tizen.org/>

para se determinar quais modificações deveriam ser feitas. Também foram feitos estudos das bibliotecas utilizadas para processamento de imagens e reconhecimento de texto, para ver como seriam feitas as modificações.

Para realizar a tarefa que lhe foi dada, o aluno escolheu que a melhor forma seria a separação do vídeo em quadros individuais que então seriam passados pelo código de reconhecimento de texto, uma vez que a biblioteca usada para reconhecimento de texto trabalha com imagens apenas e não vídeos. O vídeo então é reconstruído para ser exibido em um dispositivo eletrônico após passar por esse estágio de processamento.

Os testes para verificar o funcionamento do código são feitos em um ambiente simulando um microcontrolador. Todos os códigos foram desenvolvidos em C/C++.

## **Resultados**

### ***1) Testes de código para compressão de imagem***

Após meses de testes intensivos, o grupo no qual o aluno estava inserido foi capaz de entregar resultados para a reunião do JPEG-Pleno, onde eles foram apresentados para que o codificador a ser desenvolvido disputasse com outros possíveis candidatos para se tornar no futuro o padrão na codificação de *light-fields*.

Os resultados obtidos comprovaram a efetividade do codificador especialmente para *bitrates* mais baixos, fazendo com que os membros do comitê o aprovassem como um dos codificadores a serem usados para o futuro padrão internacional para os novos formatos de imagem.

### ***2) Desenvolvimento de códigos para reconhecimento de texto***

O código desenvolvido pelo aluno fazendo uso da biblioteca aberta de visão computacional *OpenCV*<sup>3</sup> conseguiu realizar apropriadamente o reconhecimento de texto em vídeos. Em função de o código já ter passado por processos de otimização pelo grupo que estava trabalhando no projeto antes de o aluno chegar, a eficiência, velocidade e precisão com que ocorreu o reconhecimento dos quadros individuais do vídeo utilizado para teste estava dentro das restrições de projeto previamente estabelecidas de tempo (uma vez que este código viria a se tornar uma contribuição para a plataforma Tizen).

## **Bibliografia**

- [1] Schnabel, R., Klein, R., *Octree-based Point-Cloud Compression*, Eurographic Symposium on Point-Based Graphics, 2006.
- [2] Schuetz, M., *Potree: Rendering Large Point Clouds in Web Browsers*, Master Thesis, 2016.

---

<sup>3</sup> <https://opencv.org/>

