

Sistema de Identificação de Pessoas baseado em Visão Computacional Estereoscópica

DESCRIÇÃO DO PROJETO

INTRODUÇÃO :

Com o advento dos processadores de alta performance com múltiplas CPU e GPU, dotados de funcionalidades de multimídia e comunicação integrados, torna-se tecnicamente viável descentralizar parte do processamento matemático para a borda dos sistemas (Edge Computing/Fog Computing).

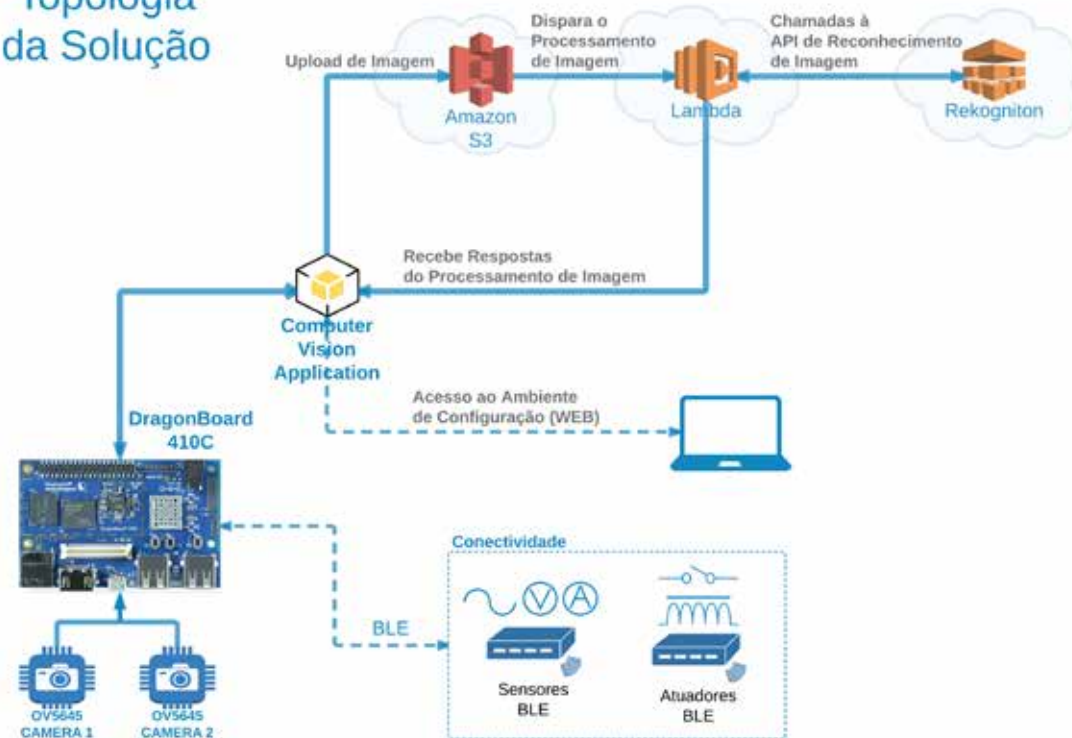
Anos atrás, o processamento de “álgebra linear” característicos dos algoritmos de “machine learning” só se viabilizava de maneira centralizada, devido aos requisitos de tamanho, consumo e custo de servidores de alta-performance. Recentemente, o custo de hardware GPU para consumidores vem caindo significativamente, ao mesmo tempo em que serviços de nuvem (Cloud Computing) especializados em aplicações de “machine learning” são abundantes.

O termo IoT (Internet das Coisas) representa hoje a combinação de diversas tecnologias antigas e originalmente batizadas de maneiras diversas: Telemetria, Automação Industrial, Wireless Data, M2M, Data-Centers, Ciência de Dados, etc.

A palavra “machine learning” também representa coletivamente técnicas maduras, dentre elas as redes neurais, que foram idealizadas há várias décadas.

Apenas recentemente, algoritmos de redes neurais ganharam notoriedade devido à alta disponibilidade de poder computacional a custo progressivamente reduzido.

Topologia da Solução



Sistema de Identificação de Pessoas baseado em Visão Computacional Estereoscópica

Cezar Menezes e Cleber Picolo

DESCRIÇÃO DO PROJETO

O PROJETO :

Acreditamos que um produto baseado na plataforma DragonBoard 410C seja viável em aplicações que usufruam da alta performance das múltiplas CPU e GPU integrado, capacidade de multimídia e comunicação integrados.

Nosso projeto pretende implementar um sistema de Visão Computacional Estereoscópica, dotado de 2 módulos de câmeras 5-megapixel com auto-foco e interfaceados à DragonBoard 410C através do padrão MIPI-CSI.

Utilizaremos BLE para a interface com sensores (ex: sensores de passagem) ou atuadores (ex: relês) BLE, para a geração de triggers de eventos ou comando de dispositivos externos.

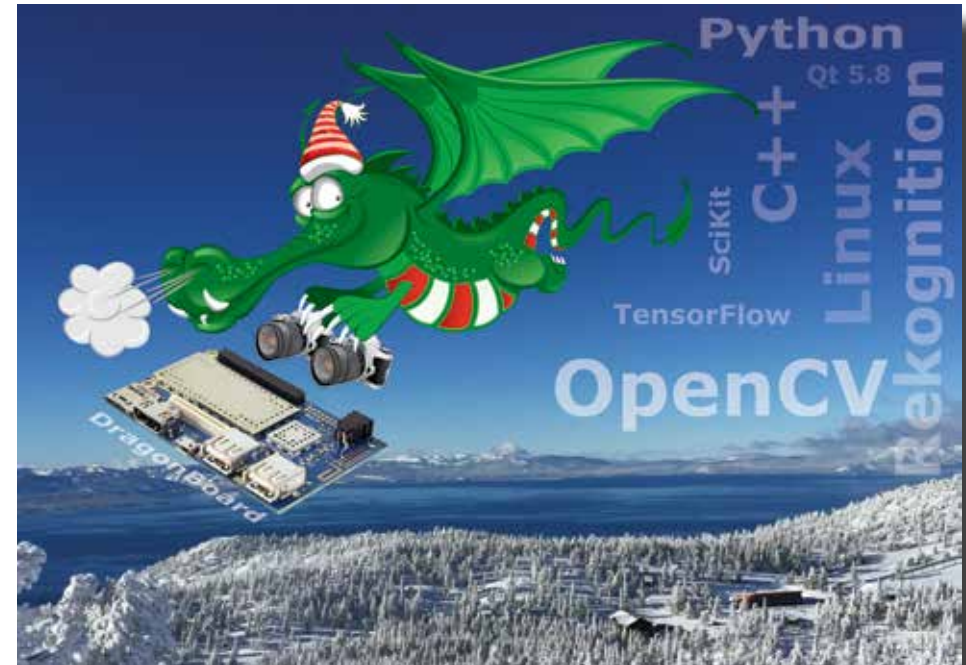
A DragonBoard 410C, graças à sua alta capacidade de processamento e multimídia, permite a utilização embarcada de stacks de software típicos em aplicações de machine learning (Python, NumPy, SciPy) e visão computacional (OpenCV, Gstreamer, V4L).

Utilizaremos duas câmeras para implementar algoritmos de geometria epipolar, visando introduzir o vetor de profundidade Z à aquisição de imagens das câmeras (vetores X,Y). Há inúmeras aplicações conhecidas para “visão estéreo” mas, neste projeto, pretendemos utilizar a informação de profundidade (distância entre objetos e o plano de imagem no eixo Z) para a geração automática de triggers de captura de imagem. Utilizaremos, também, triggers de captura de imagem provenientes de sensores BLE externos e próximos.

O processamento para detecção e extração de features (identificação

de faces a uma certa distância Z conhecida) será realizado localmente na DragonBoard410C. O processamento para reconhecimento facial será realizado em nuvem, utilizando-se API específicas de machine learning (AWS Rekognition).

Dado ao reduzido prazo do projeto, não daremos ênfase à implementação de algoritmos mais complexos de “machine learning” localmente na DragonBoard410C(ex: TensorFlow), embora acreditemos ser esta uma abordagem possível e interessante (devido à capacidade computacional disponível no processador).



Sistema de Identificação de Pessoas baseado em Visão Computacional Estereoscópica

CÂMERAS MIPI

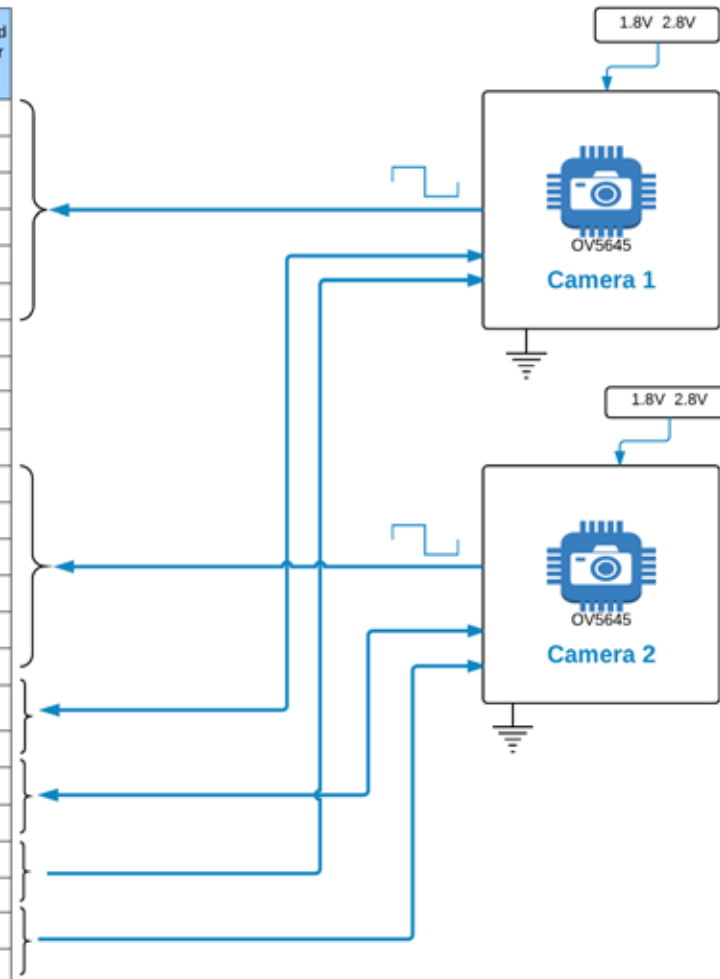
A **Dragonboard 410C** implementa duas interfaces de câmera **MIPI-CSI** :

- 1 de quadro vias (**MIPI_CSI0**)
- 1 de duas vias (**MIPI_CSI1**)

Os sinais **MIPI-CSI** são mapeados diretamente ap conector de expansão High Speed da **Dragonboard 410C**

As câmeras MIPI requerem também interface **I2C**. Para a **MIPI_CSI0**, utiliza-se a interface **I2C2**. Para a **MIPI_CSI1**, utiliza-se a interface **I2C3 (*)**

Module	SnapDragon 410E signal	96boards signal	High Speed Connector Pin	Low Speed Connector Pin
MIPI_CSI0 (primary)	MIPI_CSI0_CLK	CSI0_C+	2	
		CSI0_C-	4	
	MIPI_CSI0_D0	CSI0_D0+	8	
		CSI0_D0-	10	
	MIPI_CSI0_D1	CSI0_D1+	14	
		CSI0_D1-	16	
MIPI_CSI0_D2 (não utilizado)	CSI0_D2+	20		
MIPI_CSI0_D3 (não utilizado)	CSI0_D2-	22		
MIPI_CSI1 (secondary)	MIPI_CSI1_CLK	CSI0_D3+	26	
		CSI0_D3-	28	
	MIPI_CSI1_D0	CSI1_C+	54	
		CSI1_C-	56	
MIPI_CSI1_D1	CSI1_D0+	42		
	CSI1_D0-	44		
I2C2	I2C3_D1+	48		
	CS1_D1-	50		
I2C2	I2C2_SCL	I2C2_SCL/APQ GPIO30	32	
	I2C2_SDA	I2C2_SDA/APQ GPIO29	34	
I2C3 (*)	I2C3_SCL	I2C3_SCL/APQ GPIO15	36	
	I2C3_SDA	I2C3_SDA/APQ GPIO14	38	
PWDN0	CSI0_PWDN	APQ GPIO34		32
RESET0	CSI0_RST	APQ GPIO35		31
PWDN1	CSI1_PWDN	APQ GPIO33		34
RESET1	CSI1_RST	APQ GPIO28		33



Sistema de Identificação de Pessoas baseado em Visão Computacional Estereoscópica

Cezar Menezes e Cleber Picolo

DIFERENCIAIS DA SOLUÇÃO:

A utilização da SBC DragonBoard 410C como base de hardware na criação de um produto embarcado capaz de capturar simultaneamente imagens de duas câmeras 5-megapixel e processá-las localmente através de “stack de software” tradicional na área de conhecimento de “machine learning” e “visão computacional” coloca a solução proposta em posição vantajosa em relação a sistemas CFTV convencionais e centralizados, baseados em câmeras IP ou analógicas.

A possibilidade de acrescentar ao plano de imagem 2D (eixos X e Y) um terceiro eixo de profundidade (eixo Z) torna a solução um sensor de imagem inteligente, capaz de não apenas reconhecer padrão de objetos das cenas mas, também, de determinar a profundidade dos objetos (ou pixels) em relação ao plano de imagem. Acreditamos que muitos projetos futuros poderão se beneficiar desta solução, utilizando-a como um "sensor inteligente de captura de imagem" capaz de extrair informação de objetos ("coisas") ou pessoas a partir de imagens estéreo.

Da análise preliminar dos datasheets e “reference designs” disponíveis a respeito da DragonBoard 410C e seus acessórios, concluímos ser possível e viável o projeto e fabricação nacional de uma PCB “Mezzanine” simples e específica para o interfaceamento dos módulos de câmeras MIPI. Os componentes mais caros desse “Mezzanine” seriam os módulos de câmera (OV5645). Os demais componentes seriam, basicamente, conectores dos barramentos high/low speed da DragonBoard 410C, conectores dos módulos de câmeras, além daqueles dedicados à conversão de nível de tensão DC (1.8V e 2.8V). A nacionalização proporcionaria uma redução significativa de custo, em comparação ao custo de importação de soluções prontas.

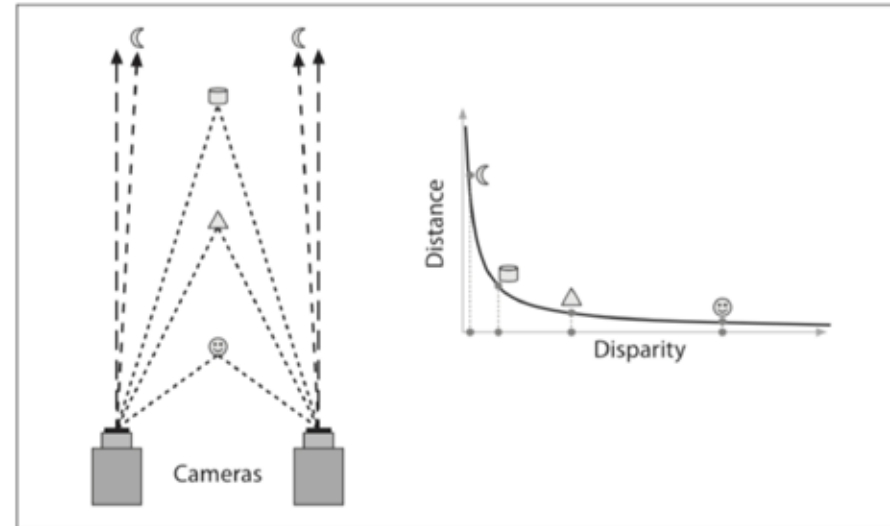
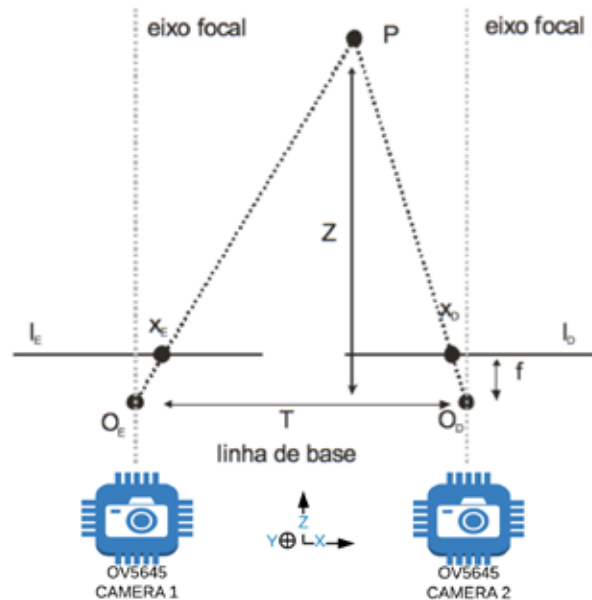
Os periféricos de comunicação Wireless (WiFi, Bluetooth e BLE) permitam não apenas a conexão à internet mas, também, a dispositivos externos (sensores BLE, atuadores BLE). Utilizar sensores sem fio (via BLE) é vantajoso em relação ao uso de cabos (isolamento elétrico, custo de instalação e conveniência).

Atualmente, a disponibilidade das APIs de “machine learning” por diferentes empresas de “cloud computing” é abundante. A concorrência neste mercado estimula não apenas o aprimoramento contínuo das funcionalidades dos serviços (APIs), mas também a redução dos custos. Neste projeto, pretendemos utilizar o serviço AWS Rekognition oferecido pela Amazon para a implementação de um sistema de reconhecimento facial de pessoas.

Alguns possíveis mercados de aplicação da solução são listados abaixo:

- **Shoppings** : Contagem de pessoas, análise de fluxo e identificação de satisfação de clientes em shoppings;
- **Segurança predial** : Controle de acesso e rastreamento de pessoas em prédios e/ou empresas;
- **Aplicações veiculares** : Identificação de motorista e passageiros, monitoramento de segurança (uso de cinto de segurança, detecção de sonolência do motorista);
- **Person of Interest** : Busca automática de suspeitos ou pessoas desaparecidas em prédios com grande movimento;
- **Identificação de clientes fidelizados** : Sistema de reconhecimento e publicidade direcionada (Ad Tech);
- **Banking/ATM** : Detecção de fraude, autenticação de usuário em frente ao ATM usando biometria facial.

Geometria Epipolar



x_E = posição no eixo X de um determinado objeto na **Câmera 1** (Esquerda)
 x_D = posição no eixo X de um determinado objeto na **Câmera 2** (Direita)

$$\text{disparidade} = (x_D - x_E)$$

A profundidade Z é inversamente proporcional à disparidade

Logo, cada ponto (*pixel*(x,y)) da imagem é composto por :

- componente cor RGB : $\text{color}(x,y) = [R(x,y), G(x,y), B(x,y)]$
- componente profundidade : $z(x,y)$

$$\frac{T}{Z} = \frac{T + x_E - x_D}{Z - f}$$

Resolvendo para Z,

$$Z = f \frac{T}{x_D - x_E}$$

Sistema de Identificação de Pessoas baseado em Visão Computacional Estereoscópica

Cezar Menezes e Cleber Picolo

PONTOS CRÍTICOS :

Na execução do projeto, planejamos dosar design de hardware e software de maneira equilibrada. Entendemos que o prazo é curto para atingirmos o estágio de fabricação de uma PCB (“mezzanine”) customizada para nossa aplicação. Mesmo assim, dedicaremos parte do trabalho ao design eletrônico relacionado à interface MIPI-CSI das câmeras.

Prevemos, também, que o “bring-up” do hardware e sistema operacional (Linaro Debian) nos demandará algum tempo. Como utilizaremos referências de design (“mezzanines”) disponíveis no mercado, esperamos encontrar informação de suporte (documentação e forums), além do eventual suporte dos fabricantes desses acessórios.

O kit de sensores (“Linker mezzanine”) poderá ser muito bem aproveitado ao longo da execução do projeto. Novas ideias surgem continuamente ao longo dos ciclos de projeto.

Após a validação do hardware (interfaceamento das câmeras e BLE) e instalação dos “stacks de software” (Sistema Operacional, Python, OpenCV, bibliotecas, etc), priorizaremos o design dos algoritmos de “geometria epipolar” (câmeras estéreo), de processamento de imagem (identificação facial) e de reconhecimento facial.

Nossa aplicação “demo” focará em “identificação de pessoas”, mas o resultado poderá ser futuramente aplicado a novos projetos que demandem “visão computacional embarcada”.

INSUMOS NECESSÁRIOS:

- 1 DragonBoard 410C (Arrow)
- 1 MIPI Adapter Mezzanine for DragonBoard 410C (AISTARVISION)
- 2 módulos de câmera 5-megapixel auto-foco com interface MIPI OV5645 (OmniVision)
- 1 Display Mezzanine Kit (LinkSprite)
- 1 Módulo Bluetooth Low Energy 4.2 RN4871 (Microchip)
- 1 Linker mezzanine card starter kit (LinkSprite)
- Acesso a internet seja via Wi-Fi ou qualquer outro meio
- 1 conta AWS para acesso aos serviços S3 e Rekognition