TP5 – Um jogo simples de encestar

Nuno Carmo Vivas – 74799 – a74799@ualg.pt

Relatório Técnico – LESTI – Computação Visual

Resumo

O objetivo deste trabalho é apresentar, em forma de jogo, as várias tecnologias aprendidas no decorrer da cadeira de Computação Visual. Estabeleceram-se metas iniciais, em forma de documento-proposta, as quais foram enviadas ao Professor e posteriormente, aprovadas. De forma incremental, foram-se completando os desafios no código e o resultado final é um jogo funcional. No final é percetível que o modelo de computação visual *MediaPipe* está bastante avançado e consegue, de forma mais ou menos estável, apresentar uma *precision* bastante alta no que se refere a identificação de *features* do corpo humano. Usaram-se dois processamentos diferentes, um para processar apenas mãos e outro para processar cara, pois foi estabelecido como objetivo pelo Professor fazer identificação da cara também. É a melhor solução para este tipo de trabalho, não obstante *Python* não suportar nativamente *“event driven development”*

# Introdução

O problema apresentado resume-se em conseguir demonstrar, de forma mais ou menos simples, o uso de técnicas aprendidas no decorrer da disciplina de Computação Visual. Estas técnicas focam-se, como o próprio nome da disciplina indica, no processamento de imagens usando, sobretudo, técnicas de Inteligência Artificial, mas também outras *libraries*, como o openCV, de modo a fazer um número grande de operações e transformações sobre as imagens para conseguir determinado objectivo. A introdução fornece o contexto para o estudo. Define a questão de investigação, a sua relevância e como o estudo pretende abordá-la. Deve atrair o interesse do leitor, estabelecendo o cenário para o que virá a seguir.

# Revisão da Literatura

Durante o semestre, procurou-se percorrer um pouco a história da computação visual, começando-se por usar modelos mais antigos, para terminar em modelos já recentes e com muito maior capacidade. Esta evolução é notória neste trabalho o qual seria impossível de realizar, pelo menos de forma pratica, usando modelos com 10 ou mais anos. Em termos práticos, se fosse opção teria optado pelo uso de apenas um modelo (e assim, um único processamento) para realizar este trabalho. Deixando de fora reconhecimentos de faces e outro tipo de processamentos. Permitiu-me perceber estas “*frameworks*” de uma forma, pelo menos, introdutória o que facilitará não só o planeamento como eventual trabalho futuro na área, caso surja.

# Materiais e Métodos

Esta secção descreve os métodos e técnicas utilizados para recolher, analisar e interpretar os dados. Deve ser detalhada o suficiente para que outros possam replicar o estudo. Deve incluir informações sobre o design do estudo, a seleção de participantes, os instrumentos e procedimentos utilizados, os métodos de análise dos dados, etc.

## Metodologia

A metodologia usada em programação foi do tipo incremental, pretendendo dividir os problemas em problemas mais pequenos e resolvendo-os um a um , de forma intercalada. Inicialmente foi definido um roteiro, discutido com o Professor e este serviu como base para a execução do trabalho. Sempre que possível, foram usadas funções para aplicar as mesmas soluções em várias zonas do código. O projeto teve 3 fases, a saber:

## Equipamento e ferramentas utilizadas

Este projeto foi inteiramente desenvolvido utilizando um Apple M1 como computador de trabalho. Como ferramentas, foram utilizadas o *Blender* para criação de imagens em três dimensões, *Visual Studio Code* como IDE, *Gimp* para processamento de imagem, *chatGPT* (versão 3.5) para produção de código, Microsoft Word como processador de texto e Github como repositório. Este repositório é publico e pode ser consultado por qualquer pessoa.

# Resultados e Discussão

Os objetivos foram conseguidos e o projeto implementa um jogo de “basquetebol virtual” onde o jogador apanha numa bola, colocada ao acaso no ecrã, com qualquer uma das mãos e a transporta para um cesto, este também, colocado ao acaso, no ecrã.

Os resultados são apresentados nesta secção em forma de texto, tabelas, gráficos ou figuras. A análise de resultados/discussão interpreta esses resultados em relação ao tema da investigação. Deve-se destacar quaisquer padrões, tendências ou relações observadas nos dados e comparar os resultados com estudos anteriores.

## Resultados Experimentais

Por aqui a tabela proposta e depois falar dos obstáculos encontrados e de as escolhas tomadas.

Por também os problemas anotados no código.

O primeiro grande obstáculo foi a definição de *pinch*. Tentou-se primeiro fazer de forma ótima, isto é, utilizando apenas o dedo indicador e o dedo gordo, mas provou-se não ser suficiente. Depois de analisado e discutido com o Professor, chegou-se à conclusão que o melhor seria fazer com todos os dedos. Depois deste código implementado, o *pinch* ficou significativamente mais robusto. Depois foi o movimento da bola. É fundamental que o deslocar da bola seja natural para o jogador. Aqui, a opção tomada foi de ter como referência apenas o dedo indicador (pela simples razão de ser aquele que tem a maior probabilidade de ser identificado). Esta opção, mostrou-se bastante eficaz e tornou o jogo bastante fluido. A bola a girar, foi um desafio superado logo na definição. Têm-se uma série de imagens em sequência, pré-carregadas numa lista, as quais são mostradas na imagem consoante o índice de frame atual (baseado na média por segundo). O resultado é um movimento fluido, mesmo em computadores como o Mac M1 que são notoriamente mais lentos (por enquanto) a usar livrarias como o openCV. Foi implementada uma função que, de forma aleatória faz *spawn* de uma caixa. Esta caixa pode ser usada para mostrar tanto a bola de basquetebol, como o cesto. O aparecer da bola em sítios ao acaso, torna o jogo mais interessante. São feitos alguns displays no ecrã. É mostrado o resultado e dificuldade no canto inferior direito. No canto superior direito a aferição dos *pinches* e na mão, os pontos actualmente detetados. Esta ultima opção foi tomada porque se verificou que havia alturas em que o algoritmo não deteta ponto nenhum e assim proporciona-se ao jogador uma pista visual do resultado dos seus movimentos. Por fim, como medida para evitar medidas inválidas no programa, foi definido que sempre que a bola sai do ecrã, é reiniciada para um ponto ao acaso no ecrã.

As palavras-chave aqui são precisão e clareza. Vale (como sempre) o bom senso. Não é necessário descrever textualmente a execução do programa linha a linha! Incluir todos os testes efetuados, recorrendo, caso seja possível, a tabelas dos resultados experimentais.

No caso de relatórios de experiências didáticas, é **particularmente grave copiar o roteiro** e deixar de observar as pequenas alterações que são sempre necessárias quando o procedimento sugerido é efetivamente posto em prática.

Devem ser incluídas as medidas e observações que servirão de base para as análises e conclusões do relatório. Mais uma vez, clareza e precisão são as palavras-chave. Tendo isto em mente, um resultado experimental similar ao apresentado na Tabela 1 deve ser apresentado no formato de tabela. Note que se evita ao máximo a utilização de linhas verticais (ver Fig. 1).

**Tabela 1**

*Contador Johnson de 4 estágios*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Pulso (*clock*) | Saída | | | |
| Q1 | Q2 | Q3 | Q4 |
| – | 0 | 0 | 0 | 0 |
| 1 | 1 | 0 | 0 | 0 |
| 2 | 1 | 1 | 0 | 0 |

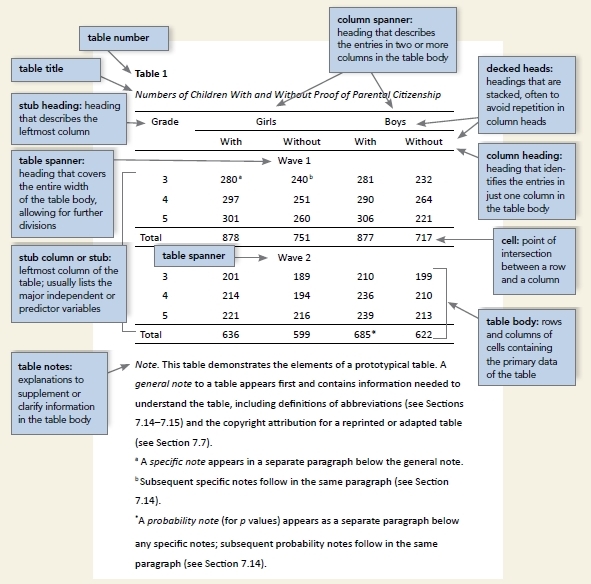


Fig. 1 – Diagrama de elementos de uma tabela, segundo a norma APA. Retirado de [Table setup (apa.org)](https://apastyle.apa.org/style-grammar-guidelines/tables-figures/tables).

No lugar de uma enfadonha listagem de métodos e números, temos uma apresentação clara e precisa, que facilita comparações e análises. **A apresentação em tabelas é sempre preferível à descrição textual.**

Observar que a legenda deve explicitar claramente o conteúdo da tabela dispensando qualquer outra explicação de metodologia. Isto nem sempre é possível, mas deve ser sempre tentado.

## Análise dos resultados

A ideia inicial – o requerimento/requirement - de ter *swipes* para alterar a dificuldade provou criar muitas dificuldades na execução já que afeta diretamente o transporte da bola de basquetebol. Depois de várias iterações, concluiu-se que a melhor hipótese seria ter reconhecimento de gestos, ou solução semelhante. Atualmente, o MediaPipe já oferece este tipo de reconhecimento, não utilizando diretamente a livraria, mas através de *tasks* e de modelos pré-treinados próprios. Este tipo de tecnologia fica fora do scope deste projeto, que visa cimentar a utilização básica das livrarias estudadas nas aulas. Se houver continuidade neste projeto esta será certamente uma opção a seguir até porque é computacionalmente mais eficiente.

Falar sobre os itens todos:

Pinch, spawn to bola, transporte da bola, spawn do cesto, scoring, dificuldades, jogo entra em pausa, e confetis,display dos scores

**É a parte mais importante do relatório**. É aqui que verificamos se o relator efetivamente está consciente dos objetivos do trabalho. Como procedimento geral, recomenda-se que os resultados de cada procedimento experimental sejam verificados. Não deve ser perdida nenhuma oportunidade de comparação entre os resultados obtidos e a teoria, ou entre os diversos procedimentos de medida. Em particular, deve-se sempre procurar valorizar os seguintes aspetos:

* Os resultados concordam com o previsto na teoria? Em que circunstâncias a concordância é melhor ou pior? Por que razões?
* No caso de comparação entre diversos procedimentos, em que circunstâncias cada um deles se destacou? É possível explicar isto a partir da teoria estudada?

A ocorrência de discrepâncias e resultados estranhos não invalida a experiência quando existirem explicações plausíveis que possam ser indicadas.

# Conclusão

Uma das conclusões a que chego ao fim deste trabalho é que o *mediapipe* é uma excelente livraria para nos iniciarmos na computação visual. Tem vários modelos pré-trabalhados, que permitem reconhecimento facial, mãos e de pose. É bastante versátil e eficiente. Não obstante, quando queremos ter mais controlo sobre o processamento e ao mesmo tempo aumentar precisão, é necessário adotar estratégias que impliquem usar modelos específicos para as tarefas que estamos a tentar executar.

O trabalho pode ser melhorado em muitos aspetos. Talvez o principal, tenha que ser um esforço suplementar para conseguir grupo de trabalho. Este trabalho foi realizado por uma única pessoa. Nestes casos, é frequente perder-se a perspetiva e estou seguro que isso influenciou algumas das decisões que foram tomadas.

Julgo que todos os objectivos propostos foram alcançados. Não obstante ainda haver caminho a percorrer, julgo que evolui bastante na linguagem de programação Python e que os conceitos de computação visual ficaram bastante bem cimentados como resultado do trabalho posto neste projecto.

# Agradecimentos

Gostaria de agradecer à minha Mulher, o apoio constante durante esta fase, diferente, da minha vida. Ao Professor Jaime, a dedicação e disponibilidades mostradas para com os alunos desta disciplina.

# Referências

Indices MediaPipe para dedos: <https://developers.google.com/mediapipe/solutions/vision/hand_landmarker>

Lista todas as fontes citadas no relatório. Deve-se seguir um estilo de citação apropriado (APA, MLA, Chicago, etc.) Ao citar uma fonte, deve-se fornecer informações suficientes para que os leitores possam localizá-la.

Esta secção **não deve ser uma lista de livros ou artigos acerca da teoria envolvida.** Devem ser indicadas referências de onde efetivamente foi retirado material de suporte teórico, ou onde se sabe estar descrito em maior detalhe um aspeto abordado superficialmente no relatório. Cada item deverá conter necessariamente: **[número da referência] Iniciais e sobrenomes de todos os autores, título, editora e local de edição, volume, páginas, ano da edição.**

# Apêndices (se necessário)

Esta secção pode incluir qualquer material adicional, que tenha sido produzido pelos autores, e que complemente as informações no relatório.

# Anexos (se necessário)

Esta secção pode incluir informações adicionais, produzidas por terceiros (que não os autores), como transcrições de entrevistas, fotografias, mapas, etc. que não foram incluídos no corpo principal do relatório.

**Recomendações para a elaboração de relatórios**

De um modo geral, um relatório bem escrito deve permitir que um leitor, com suficiente conhecimento na área em estudo:

* Compreenda de imediato os objetivos e o âmbito do trabalho.
* Recorde os principais conceitos da teoria envolvida, sendo informado sobre bibliografia onde pode obter mais informação.
* Compreenda como reproduzir o trabalho realizado, com todos os detalhes relevantes.
* Acompanhe as conclusões do autor, baseadas nos resultados obtidos.

**Objetivos de um Relatório Técnico**

Um relatório técnico serve vários propósitos essenciais na comunicação de informação técnica e científica. Os principais objetivos de um relatório técnico incluem:

## Comunicar Informações

O objetivo principal de um relatório técnico é comunicar informações específicas a um público-alvo – o que pode envolver a descrição de um processo, a explicação de um problema técnico, a apresentação de resultados de investigação ou a explicação da implementação e operação de um sistema ou equipamento técnico.

## Fornecer Análises e Interpretações

Além de apresentar informações, também pode incluir uma análise ou interpretação dessas informações, o que pode envolver a avaliação da eficácia de um processo, o exame das causas e efeitos de um problema, ou a discussão das implicações dos resultados de uma investigação.

## Documentar Procedimentos e Resultados

Os relatórios técnicos servem como um registo oficial de procedimentos, métodos e resultados. Isto é especialmente importante em ciência e engenharia, onde a replicabilidade é crucial. Por isso, os relatórios técnicos devem ser precisos e detalhados, fornecendo informações suficientes para que outros possam entender e replicar o trabalho.

## Facilitar a Tomada de Decisões

Podem ser usados ​​para apoiar a tomada de decisões, pois fornecem informações factuais e análises que podem ajudar os gestores a tomar decisões informadas. Por exemplo, um relatório técnico pode ajudar a decidir se uma nova tecnologia deve ser adotada, ou se uma abordagem específica deve ser usada para resolver um problema.

## Contribuir para o Corpo de Conhecimento

Contribuem para o corpo geral de conhecimento dentro de uma disciplina específica. Ao documentar os métodos e resultados, permitem que outros na área beneficiem do trabalho realizado, avançando assim o campo de estudo.

**Formato para os relatórios**

Papel A4. Todo o material impresso (incluindo texto, ilustrações e gráficos) deve ser mantido dentro de uma área de 17.5cm de largura e 22.5cm de altura. O número da página deve ficar 1.25cm abaixo do texto e centrado.

O cabeçalho da primeira página deve ser no formato de 1 coluna, como exemplificado acima. Centre o título e deixe um espaço de 12 pontos depois do texto. Em seguida coloque o nome do aluno, número, curso, nome da disciplina, data e deixe um espaço de 12 pontos. O documento deve ser escrito numa fonte do tipo Serif, como por exemplo a *Times New Roman*, *STIX Two Text*, *Cambria*, ou similar. O título deve usar a fonte em negrito, de 14 pontos (pt). Em seguida deixe 2 linhas em branco.

O corpo do texto deve estar no formato de 2 colunas (exceto figuras e tabelas, se for necessário; neste caso são colocadas no topo ou no fim das páginas). Divida a última página em duas colunas iguais. As colunas devem ter 8.3cm de largura.

O texto deve ter espaçamento simples (*single-spaced*) e fonte tipo Serif de 11pt. A indentação da primeira linha de cada parágrafo deve ser de 0.5cm, mas apenas em parágrafos sequenciais. Nos restantes casos não se usa indentação, como por exemplo, no primeiro parágrafo de cada secção, ou quando já exista uma separação vertical significativa entre blocos de texto. Todo o texto deve ter uma apresentação “justificada” (*justified*), e ser hifenizado automaticamente (*hyphenated*).

O título de cada secção (e.g., **1. Introdução**) deve ser em negrito com dimensão de 12pt e o de cada subsecção (e.g., **1.1. Exercício preliminar**) deve usar fonte 11pt em negrito.

O número total de páginas não deverá exceder as seis, incluindo tabelas e figuras. A listagem completa do código-fonte deve ser colocada em apêndice.