

Comparação de Marcadores Fiduciários

A estimativa de pose é essencial para várias aplicações da área de visão computacional. No caso do UAV-ART, esta estimativa será necessária para reconhecer a posição e ângulos da aeronave em relação à pista, de modo a realizar a aproximação e o alinhamento desta durante a aterragem. O processo de estimativa de pose baseia-se em encontrar correspondências entre pontos do entorno real e as suas projeções na imagem captada pela câmara. [1] Usualmente este passo é difícil de modo a que é bastante comum o uso de marcadores fiduciários para facilitar este procedimento, especialmente na área da robótica. O principal benefício destes marcadores é que um único marcador fornece suficientes correspondências (normalmente os seus quatro cantos) para obter a posição da câmara. Além disso, a codificação interna binária torna-os robustos de forma a permitir a possibilidade de aplicar técnicas de deteção e correção de erro.

Deste modo a equipa de visão decidiu investigar mais a fundo vários tipos de marcadores fiduciários, analisando as suas vantagens e desvantagens, para que no fim seja escolhido o marcador que melhor se enquadrar às necessidades do projeto. Neste caso, é necessário um marcador que seja detetado a grandes distâncias, com uma computação relativamente baixa de forma a não interferir com o funcionamento de outros algoritmos essenciais (por exemplo, o Autopilot), e cujas estimativas de distância e atitude apresentem uma boa exatidão.

Neste relatório comparar-se-ão três tipos de marcadores fiduciários: os "Aruco Markers", os "STag Markers" e marcadores próprios desenhados pela equipa, cujas utilidades serão postas em evidência. Inicialmente será feita uma breve introdução e caracterização de cada marcador, de seguida explicar-se-ão os testes realizados. Finalmente, os resultados destes ensaios serão analisados para pôr em evidência as qualidades e falhas de cada um.

1 Aruco Markers

Os Aruco Markers são um tipo de marcador fiduciário bastante popular, sendo utilizado em diversos projetos de robôs e quadrirotóres, o que se deve principalmente a sua fácil deteção. Outro aspecto que o caracteriza é o seu vasto número de bibliotecas, apresentando no OpenCV 16 bibliotecas com diferentes tamanhos e, consequentemente, um número total de códigos ID diferentes. [2]

O Aruco Marker é composto por um quadrado preto como base com uma ampla margem nos bordos cujo interior apresenta uma matriz binária que determina o seu identificador. [3]

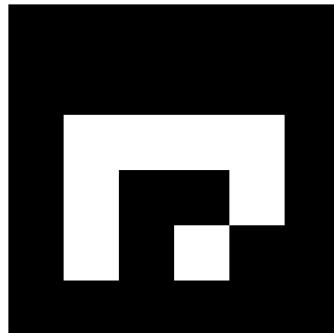
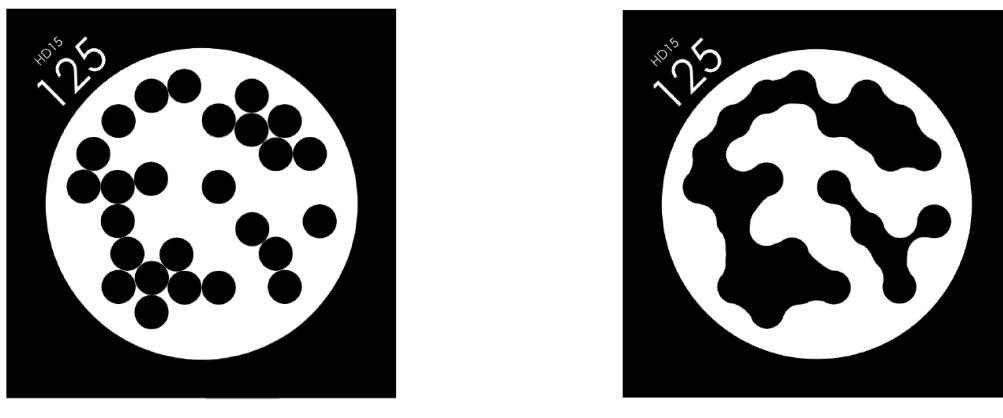


Figura 1: Marcador Aruco ID 1 da biblioteca 4x4.

2 STag Markers

Os STag Markers são um tipo de marcador fiduciário recentemente desenvolvido que se foca na estabilidade de medições da estimativa de pose e na robustez perante o movimento. [4] Uma vez que estas duas características são bastante importantes para o caso de uma câmara colocada num UAV (onde existem certas vibrações e cujas medições precisam de ser exatas), decidiu-se testar este tipo de marcadores.

A estrutura deste marcador consiste num quadrado preto como base com um círculo branco no centro. Dentro do círculo encontra-se a área de codificação, onde a identificação do marcador é armazenado. A área de codificação contém representações de bits em forma de discos. De modo a que o código seja corretamente lido na presença de ligeiros erros de localização, as lacunas entre as representações de bits vizinhas são preenchidas. Isto também reduz os elementos de alta frequência no padrão, resultando num menor número de contornos detetados no marcador. [4]



(a) Codificação inicial do marcador

(b) Marcador final

Figura 2: Marcadores STag com ID 125 da biblioteca HD15.

3 Marker C

Uma das vantagens dos marcadores fiduciários como os arUco ou os STag é a possibilidade de atribuir a cada marcador um ID único. Isto permite detectar até 16 IDs diferentes para um marcador com 4x4 bits, ou 256 IDs para um com 6x6 bits. No entanto, esta informação no nosso caso não é relevante, pois nós apenas queremos detectar o local de aterrissagem do UAV.

Como tal, tentou-se criar um novo marcador que, por possuir menos informação, iria permitir alcançar maior distância de detecção poupar tempo de computação.

Começou-se por criar-se uma série de marcadores diferentes com as formas mais básicas, isto é, círculos, triângulos e quadrados por achar-se que seriam as mais fáceis de detectar.

A estrutura básica de cada marcador eram dois ou três objectos dispostos com cores opostas. Devido à necessidade de calcular a distância da câmara para o marcador, à partida marcadores em que o exterior era um quadrado seriam melhores para isso mas decidimos testar com circulares também.

Na Figura 3 pode observar-se os marcadores que decidimos testar.



(a) A-D

(b) E-G

Figura 3: Marcadores próprios A-G

Testou-se cada marcador em testes de detecção a 1, 2, 3, 4 e 5 metros. Os marcadores que tiveram melhores resultados foram os marcadores C e G. Verificou-se que os marcadores que tinham um triângulo não eram muito promissores, pois a grandes distâncias o triângulo deformava-se muito e era facilmente confundido com um quadrado ou outra forma geométrica. Também se verificou que, como seria de esperar, quantos mais elementos geométricos o marcador tinha, menor era a distância máxima de detecção. Como tal, escolheu-se continuar a estudar os marcadores C e G por duas razões:

1) foram os que obtiveram maior distância de detecção máxima

2) intuitivamente pareciam os mais promissores. Excluindo os marcadores com três elementos, sobravam o C, D e G. No entanto, o marcador C (círculo dentro de círculo) vinha com a dificuldade de calcular a distância através da função `solvePnP()` por não possuir 4 cantos.

Em teste posteriores verificou-se que o círculo do marcador G também sofria de uma grande distorção, e por isso decidiu-se investigar apenas o marcador C.

O fluxograma de detecção do marcador C pode ser observado na Figura 4:

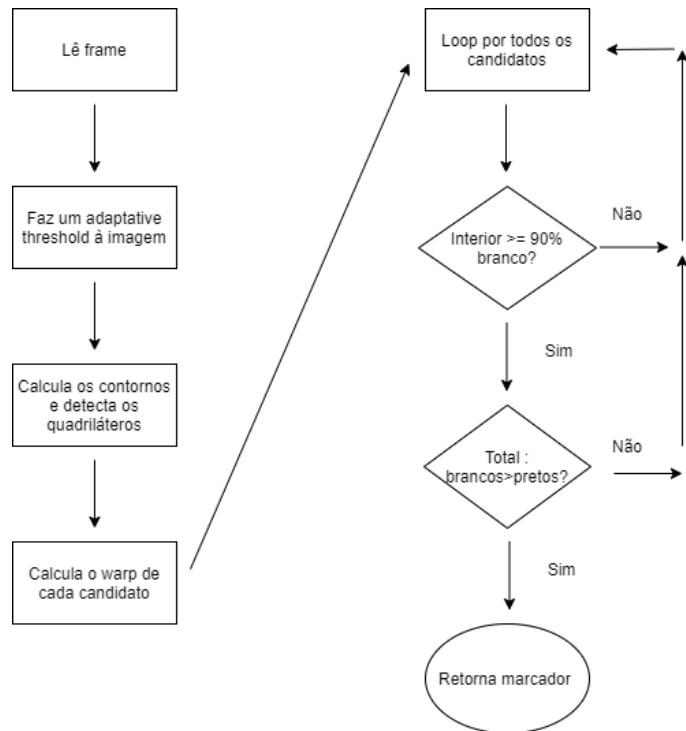


Figura 4: Fluxograma do algoritmo de detecção do marcador C

Os critérios dentro do *loop* por todos os candidatos foram escolhidos heuristicamente após diversas tentativas. Chegou-se à conclusão que apenas ver a cor do interior do marcador e a proporção total produzia os melhores resultados para o *trade-off* distância/reduzir falsos positivos. Como a grande distância a distorção do marcador é muito elevada, esta foi a maneira que verificou funcionar para a maioria dos casos.

O marcador C final pode então ser observado na Figura 5.

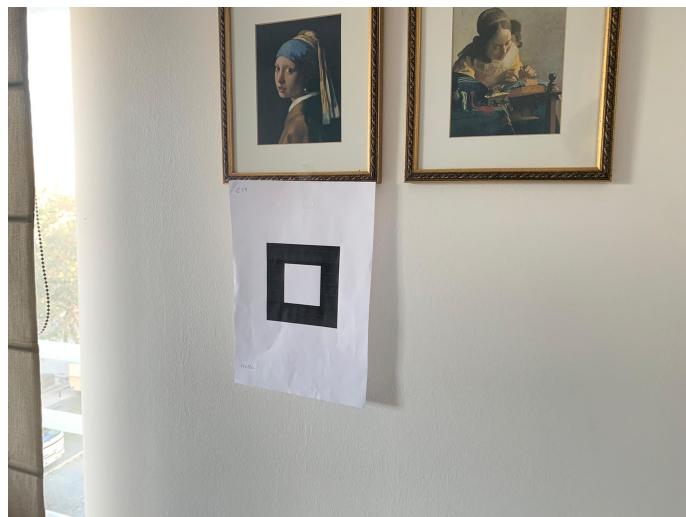


Figura 5: Marcador C

4 Resultados

Os testes realizados irão avaliar duas propriedades: a taxa de deteção e a exatidão da estimativa de distância. Para ambos testes foram gravados vários vídeos com distâncias diferentes para cada marcador.

Relativamente à taxa de deteção, em cada vídeo será contabilizado o número de frames em que o marcador foi reconhecido e este valor será dividido pelo número total de frames, obtendo assim a taxa de deteção. Esta taxa será utilizada para encontrar a distância máxima de deteção para cada marcador.

No que respeita à estimativa da distância, em cada vídeo será estimada esta variável e depois será comparada com a distância real, medida obtida por uma fita métrica. Apesar de ainda ser possível estimar a atitude de cada marcador, a exatidão desta não foi avaliada devido à falta de instrumentos que fornecessem o verdadeiro ângulo do marcador em cada prova.

É importante realçar que os resultados serão divididos em duas comparações: Aruco vs STag e Aruco vs Marker C.

4.0.1 Dataset

Os ensaios foram realizados no entorno ilustrado na Figura 6. A distância entre a câmara do computador e o centro do marcador era medida com uma fita métrica e de seguida realizava-se a gravação de dados. Os testes foram elaborados com diferentes distâncias (100, 150 e 200 cm) e para ambos tipos de marcador. O tamanho de cada marcador foi de 10cm. Os marcadores específicos utilizados são ilustrados na Figura 7



Figura 6: Montagem realizada para as comparações dos marcadores.



(a) Marcador Aruco ID 1, biblioteca Original



(b) Marcador STag ID 001, biblioteca HD15



(c) Marcador C

Figura 7: Imagem de marcadores captadas na montagem (distância de um metro).

É de salientar que o computador utilizado para realizar estes ensaios foi um *HP Pavilion Laptop 15* com um processador *AMD Ryzen 5*, uma memória RAM de 12 GB e um sistema operativo *Windows 10*. No que respeita ao software, a versão de Python utilizada foi a 3.8.4 com a biblioteca OpenCV 4.3.0. Para a deteção de marcadores STag foi utilizada a linguagem C++ com a versão 9.2.0 do compilador G++ e a versão 3.4.13 do OpenCV. Finalmente para os marcadores C foi utilizada também a linguagem C++ com a mesma versão de compilador mas uma versão do OpenCV mais recente, 4.3.0.

Outro aspeto a realçar é que durante as experiências foram utilizadas duas resoluções de imagem: uma de qualidade média (SD) com 640x480p e outra de alta definição (HD) com 1920x1080p. Por simplificação, os resultados em SD serão apresentados como "STag" e os de alta resolução como "STag HD".

4.1 Aruco VS STag

4.1.1 Detection Rate

Calculou-se então a detection rate para ambos marcadores em diferentes resoluções, obtendo assim a tabela seguinte:

Tabela 1: Percentagens de taxa de deteção

Distâncias	Aruco	STag	Aruco HD	STag HD
1m	100%	100%	100%	100%
1.5m	99.84%	98.90%	100%	100 %
2m	100%	25.78%	100%	96.84%



Os resultados obtidos para o Aruco e o STag em SD demonstram que existe uma grande divergência na taxa de deteção a partir da distância de 1.5m, chegando a uma diferença de aproximadamente 75% nos 2 metros. A partir disto confere-se que perante uma câmara de qualidade média, os Aruco apresentam uma melhor taxa de deteção que os STag.

Ao aumentar a qualidade de imagem, percebe-se que a taxa de deteção do STag aumenta significativamente, embora continue a ser superado pelo Aruco.

Para além do cálculo da detection rate para distâncias predefinidas, também foi realizado um teste onde a distância foi aumentada até atingir uma taxa de deteção por baixo dos 10 %. Esta distância pode ser vista como uma estimativa aproximada da distância máxima do marcador. Como resultado deste tipo de ensaio, observou-se que o STag apresenta a menor distância máxima com 2.2m, enquanto que o STag HD apresentou 3.2m de distância máxima, aumentando assim significativamente. Por parte dos Aruco, para a qualidade standard há uma distância de 2.8m e para HD essa distância aumenta para os 3.2m. Sendo assim, consegue-se verificar que em resoluções normais os Aruco são superiores em distância máxima de deteção. No entanto, ao aumentar a qualidade de imagem, ambos marcadores conseguem apresentar aproximadamente a mesma distância máxima.

4.1.2 Exatidão

Nos ensaios de exatidão de distância iniciou-se por comparar a diferença com uma qualidade de imagem média ambos tipos de marcadores, obtendo assim os gráficos ilustrados na Figura 8.

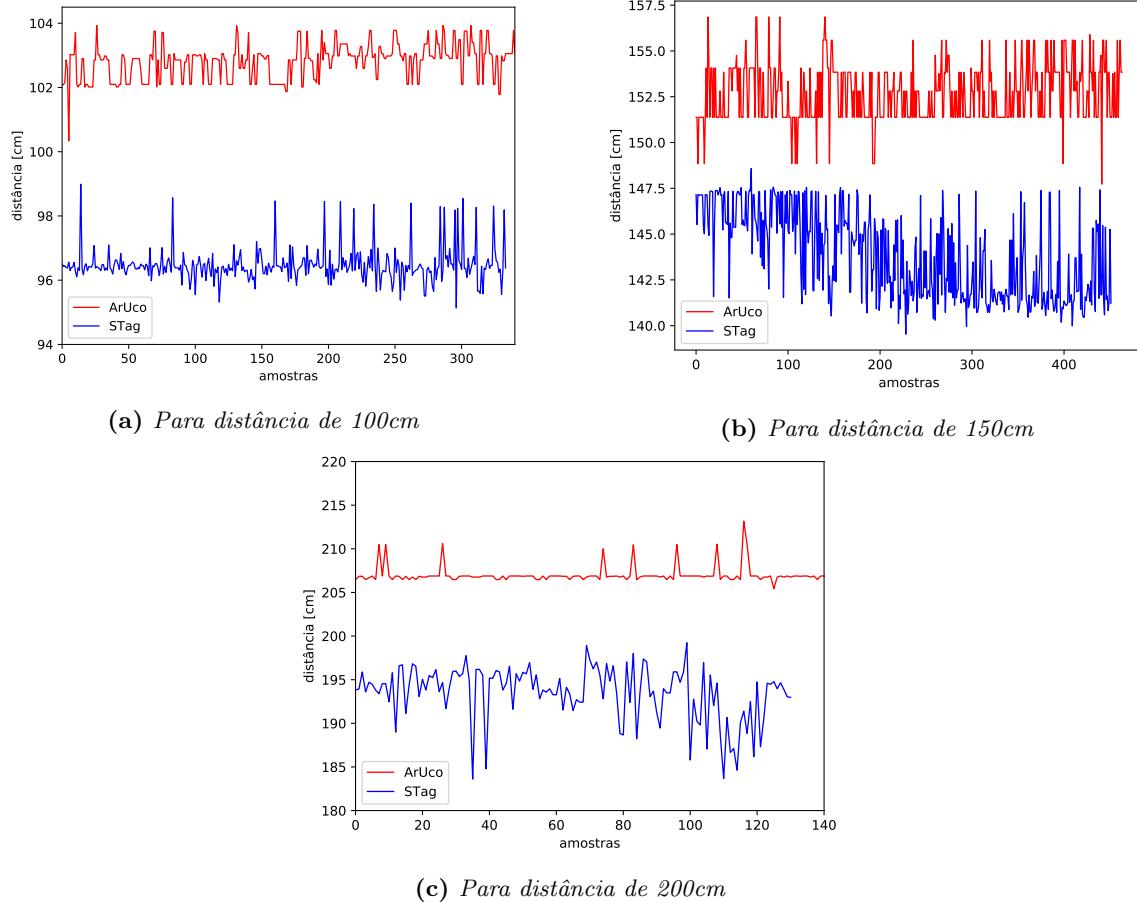


Figura 8: Gráficos de comparação das distâncias estimadas pelos marcadores durante cada amostra.

Verifica-se então à primeira vista que ambos métodos de estimação apresentam bastantes variações. Também se observa que a estimativa dos Aruco acaba por estar ligeiramente mais perto do valor real. Fazendo novamente os ensaios com uma alta resolução, obtiveram-se os resultados da Figura 9.

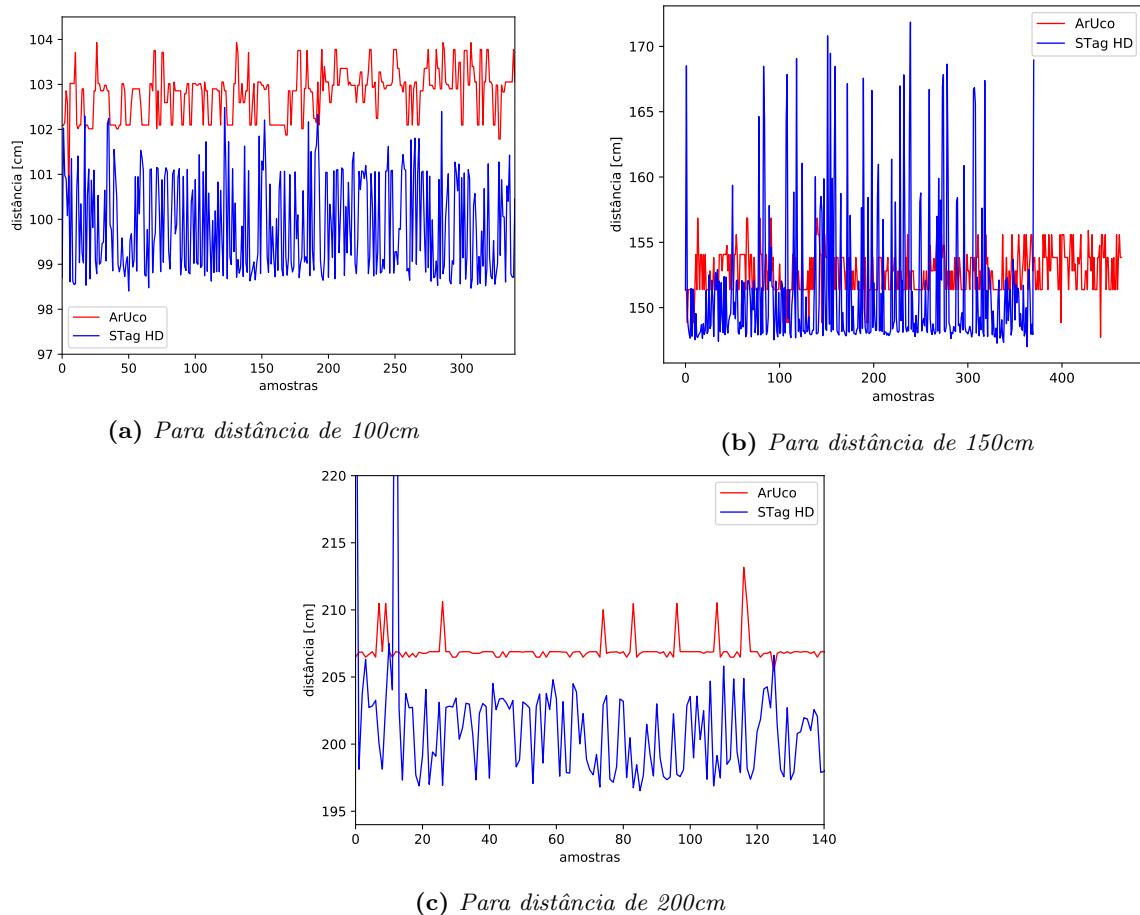


Figura 9: Gráficos de comparação das distâncias estimadas pelos marcadores durante cada amostra.

Pelo gráfico 9 infere-se que existe uma variação de valores ainda maior do que anteriormente na Figura 8 para os STag markers. Mesmo assim, observa-se que os valores de medição deste oscilam mais perto do valor real. De modo a analisar melhor estes dados calcularam-se alguns parâmetros como a média (μ), a variância (σ^2), o erro absoluto médio (ε_{abs}) e o erro relativo médio (ε_{rel}). Estes parâmetros são importantes para analisar e comparar as medições dado que os sinais de medição apresentam um comportamento aleatório.

Tabela 2: Parâmetros estatísticos de dados adquiridos durante as medições

	d = 100cm				d = 150cm				d = 200cm			
	μ	σ^2	ε_{abs}	$\varepsilon_{rel} [\%]$	μ	σ^2	ε_{abs}	$\varepsilon_{rel} [\%]$	μ	σ^2	ε_{abs}	$\varepsilon_{rel} [\%]$
Aruco	103.05	0.33	3.05	3.05	152.68	2.54	2.68	1.79	208.35	4.16	8.36	4.18
STag	96.47	0.28	3.53	3.53	143.81	5.66	6.19	4.13	193.50	9.95	6.50	3.25
STag HD	99.86	1.20	0.14	0.14	151.33	29.15	1.33	0.88	201.34	17.03	1.34	0.67

É de salientar que os resultados dos Aruco em HD não tiveram mudanças significativas aos de SD, logo para evitar redundâncias não foram colocados os resultados deste na Tabela 2. Contrariamente, para os STag, observa-se que ao aumentar a resolução o erro de estimativa diminui bastante, chegando a ser menor do que os Aruco, apresentando um erro relativo máximo de 0.88%. Apesar de ter a melhor exatidão, a variância do STag HD acaba por ser a pior respetivamente à dispersão dos seus valores (maior variância).

4.1.3 Tempo de Processamento

Durante as provas anteriores manifestou-se um fator que foi de extrema relevância para o estudo: o tempo de processamento. Verificou-se que existe uma diferença bastante significativa entre os Aruco e os STag, obtendo uma frequência de 15 FPS e 1 FPS, respetivamente (para uma câmara que filma aos 30 FPS). Isto demonstra uma alta carga computacional, especialmente por parte dos STag Markers. É de salientar que estes resultados foram obtidos com as especificações mencionadas na secção 4.0.1

4.2 Aruco VS Marker C

4.2.1 Detection Rate

Calculou-se novamente a detection rate para ambos marcadores em diferentes resoluções, obtendo assim a tabela seguinte:

Tabela 3: Percentagens de taxa de deteção

Distâncias	Aruco	Marker C	Aruco HD	Marker C HD
1m	100%	100%	100%	100%
1.5m	99.84%	99.31%	100%	100%
2m	100%	96.99%	100%	100%

Os resultados obtidos demonstram uma pequena divergência entre a taxa de deteção dos Aruco e o Marcador C. Em SD, a taxa de deteção do próprio marcador é um pouco mais baixa do que o Aruco. Apesar disso, ao aumentar a qualidade de imagem nota-se que esta diferença é compensada.

Tal como na comparação anterior, observou-se a estimativa da distância máxima dos marcadores. Deste modo, verificou-se que o marcador C apresenta uma distância máxima de 3.0m enquanto que os Aruco possuem uma distância maior de 3.2m em HD. Ao diminuir a qualidade, confere-se que as distâncias máximas dos Aruco e os Marker C diminuem para 2.8m e 2.7m, respetivamente.

4.2.2 Exatidão

Tal como foi elaborado na secção 4.1.2, compararam-se as estimativas de distância dos marcadores em qualidade SD na Figura 10.

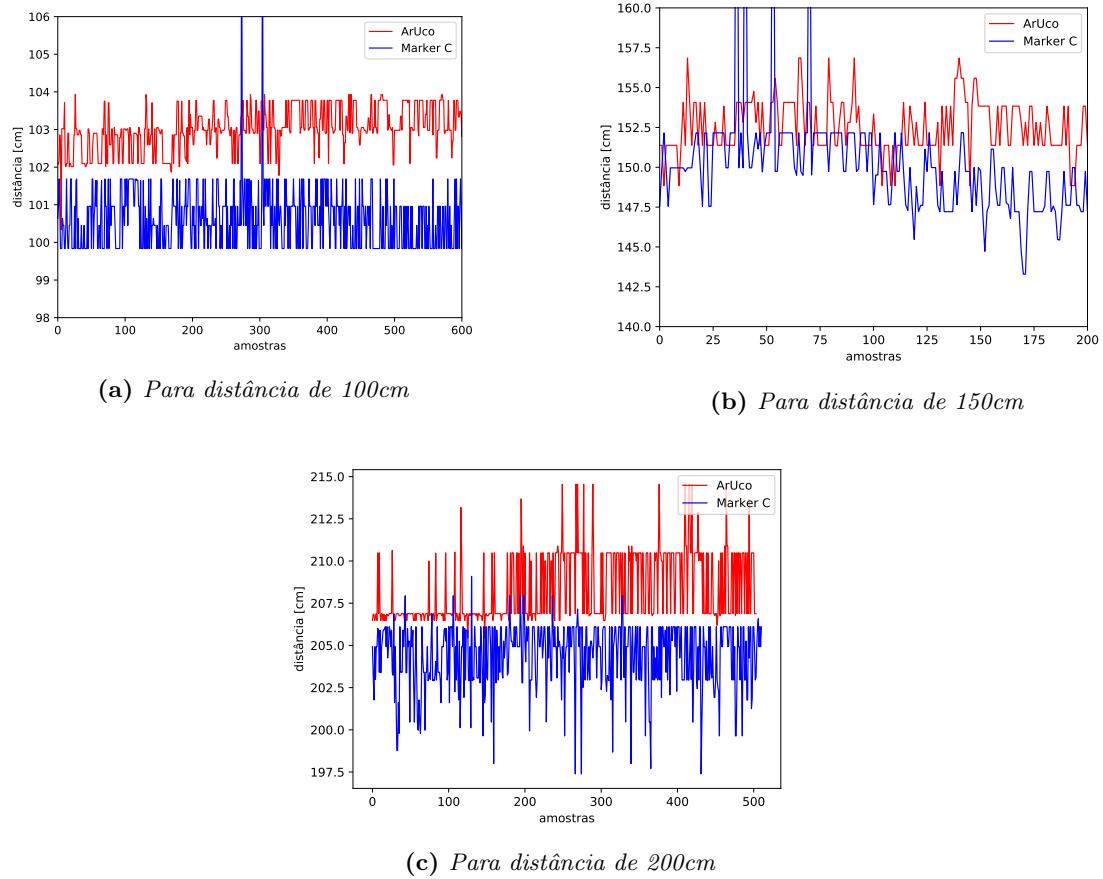


Figura 10: Gráficos de comparação das distâncias estimadas pelos marcadores durante cada amostra.

À primeira vista nota-se que o marcador C encontra-se mais perto do valor real do que os Aruco. De modo a confirmar isto e analisar os dados de uma forma mais objetiva, calcularam-se os parâmetros de μ , σ^2 , ϵ_{abs} e ϵ_{rel} .

Tabela 4: Parâmetros estatísticos de dados adquiridos durante as medições

	d = 100cm				d = 150cm				d = 200cm			
	μ	σ^2	ϵ_{abs}	ϵ_{rel} [%]	μ	σ^2	ϵ_{abs}	ϵ_{rel} [%]	μ	σ^2	ϵ_{abs}	ϵ_{rel} [%]
Aruco	103.05	0.33	3.05	3.05	152.68	2.54	2.68	1.79	208.35	4.16	8.36	4.18
Marker C	100.57	0.91	0.57	0.57	150.22	18.43	0.22	0.15	204.35	3.79	4.35	2.18

É de notar que durante esta análise não foi abordado em detalhe as medições feitas em HD, uma vez que não foi verificado uma diferença significativa entre os resultados em SD e os de HD. Relativamente aos parâmetros obtidos na Tabela 4 confirma-se o que foi observado na Figura 10, ou seja, uma maior exatidão por parte do marcador C. Verifica-se então que o marcador C apresenta erros relativos menores do que o Aruco para todos os ensaios realizados. Também se verifica que os Aruco apresentam uma menor variância no geral, embora esta diferença não seja muito significativa. Depara-se também na Figura que nas medições existem alguns picos, mas estes podem ser desprezados utilizando um filtro mediano.

4.2.3 Tempo de Processamento

Para os testes com o marcador C, os tempos de processamento estão expostos na Tabela 5. Estes ensaios foram feitos com um processador i7-88550U CPU @ 1.80GHz x 8. A versão Python é 3.8.1 e a de C++ foi executada com um GCC 5.4.0.

O cálculo do tempo de processamento foi feito com os vídeos "*C_fast.MOV*" e "*AR_fast.MOV*". As *FPS* foram calculadas fazendo uma média do tempo que o programa demorou a processar cada *frame*. No vídeo real, as *FPS* podem alterar consoante os objectos a processar em cada *frame*.

Tabela 5: Tempo de processamento (FPS)

Marcador	Python	C++	Named Pipe
Aruco	29.105	-	-
Marker C	6.865	21.267	18.765

Como podemos ver, a biblioteca dos Aruco está bastante optimizada, o que não é uma surpresa. A biblioteca está escrita em C++, por isso não achámos relevante testar o tempo de processamento para C++.

O marcador C tinha resultados bastante fracos em Python. Para contornar isso, começou-se por refazer o código para C++, com bastante sucesso como se pode ver pela tabela. No entanto, como o resto da equipa UAV usa o Python para fazer a comunicação entre equipas, e mais importante, o código da equipa de Controlo está escrito em Python. Daí a necessidade de, apesar do grosso do processamento ser em C++ devido à sua rapidez, tivemos que fazer a comunicação do C++ para o Python através de um pipe entre os dois programas. Apesar do tempo de processamento final estar um pouco longe dos Aruco, consegui-se uma grande melhoria.

4.3 Conclusões

Como podemos verificar nos resultados, na maioria das situações, os Aruco Markers e os STag Markers apresentam estimativas satisfatórias para a medição de distâncias, apresentando erros relativos máximos menores do que 5%. Mesmo assim, ainda há alguns aspetos a melhorar nas medições de ambos, como a elevada dispersão dos valores calculados, especialmente a longas distâncias. No futuro será necessário implementar algum método para estabilizar os dados fornecidos pela câmara (ex: um Extended Kalman Filter).

Também se conferiu que ao aumentar a resolução da câmara, é possível obter melhores resultados por parte dos STag superando a exatidão dos Aruco. Contudo, nestes ensaios observou-se que a maior desvantagem dos STag é o seu tempo de processamento, que é muito superior aos do seu marcador concorrente, acabando por não compensar o seu aumento de precisão. Podemos então afirmar que a escolha entre os Aruco e o STag dependerá principalmente da preferência entre rapidez e exatidão. Em relação ao *Detection Rate*, este parâmetro acaba por ser bastante parecido quando a resolução da câmara é alta, obtendo também uma distância máxima de deteção parecida. Não obstante, quando esta baixa, os Aruco acabam por superar os STag de forma abismal.

Em suma, os STag acabam por apresentar a vantagem da exatidão mas a sua deteção pouco otimizada (uma vez que é um marcador mais recente) e as suas medições dispersas dão muito a desejar ainda. Sendo assim, optou-se por analisar e testar outra alternativa, o marcador C.

A partir dos resultados obtidos verificou-se que o marcador C também apresenta estimativas de medição



bastante boas, com um erro relativo abaixo dos 3%. Pelo que nestes ensaios o marcador C acabou por ser mais exato do que o Aruco. Não obstante, ainda existe o mesmo problema dos marcadores anteriores, a sua elevada dispersão nos valores estimados. Relativamente à estimativa da distância de deteção máxima, observou-se nos resultados obtidos que quer para SD, quer para HD, a distância máxima dos Aruco é ligeiramente superior à do marcador C, mesmo assim ambas são semelhantes.

No que respeita ao tempo de processamento, conseguiu-se melhorar a performance do algoritmo de deteção do marcador C implementando o código em C++ em vez de Python. Mesmo assim, o tempo de processamento continua ser maior no marcador C, visto que o código dos Aruco está bastante otimizado. Logo, verifica-se que o Aruco supera o nosso marcador em velocidade.

Desta forma, pode-se constatar que nos ensaios realizados os Aruco foram melhores no que respeita à distância máxima de deteção e tempo de processamento, enquanto que o marcador C destacou-se pela sua exatidão nas medições. Não obstante, é preciso ter em consideração que estas provas são uma pequena amostra de comparação entre os marcadores, de modo a que, para obter conclusões mais gerais seria necessário realizar mais ensaios, adquirindo assim mais dados para analisar. Mesmo assim este relatório serve para analisar e comparar os marcadores de uma forma superficial.

Futuramente será necessário realizar mais provas aos marcadores para testar a sua robustez perante o movimento, o efeito da oclusão e a variação das distâncias limite em função da dimensões de cada marcador.

Equipa de Visão,



Santiago Rodriguez, Colaborador
Contacto: santiago.n.rodriguez@tecnico.ulisboa.pt.



Vasco Araújo, Colaborador
Contacto: vasco.faden.araujo@tecnico.ulisboa.pt.

2021/05/21

Referências

- [1] R. Szeliski. *Computer Vision: Algorithms and Applications*. London:Springer-Verlag, UK, 2010.
- [2] OpenCV Library,
<http://opencv.org> Acedido em: 18/02/2021
- [3] S. Garrido-Jurado, R. Muñoz-Salinas, F.J. Madrid-Cuevas, M.J. Marín-Jiménez. *Automatic generation and detection of highly reliable fiducial markers under occlusion*. Córdoba University 2014
- [4] Burak Benligiray, Cihan Topal, Cuneyt Akinlar. *STag: A Stable Fiducial Marker System*. 2019