



Mestrado em Engenharia Eletrotécnica

Sistemas de Monitorização e Proteção baseados em Visão 3D - Desenvolvimento de uma Aplicação de Segurança e Proteção Industrial utilizando Sensores RGB-D

Trabalho de Projeto apresentado para a obtenção do grau de Mestre em Engenharia Eletrotécnica – Área de Especialização em Automação e Comunicações em Sistemas Industriais

Autor

Márcio Filipe Santos Marques

Orientador

Doutor Fernando José Pimentel Lopes

Professor do Departamento de Engenharia Eletrotécnica
Instituto Superior de Engenharia de Coimbra

Coimbra, outubro, 2017

Dedico, com imensa honra e gratidão, o subseqüente Trabalho de Projeto à minha família e diletos amigos, reconhecendo todo o apoio facultado no decorrer deste percurso.

Aos meus Pais, Maria Manuela e Albino Marques, ao meu irmão Sílvio Marques, à minha cunhada Sandra Fernandes, à minha afillhada Sol Marques e ao meu sobrinho Saúl Marques, pelo amor, apoio, confiança e motivação incondicional, que sempre me encorajam para a conquista dos meus desafios e para a superação das diversas dificuldades ao longo da vida.

*“Nunca ande pelo caminho traçado, pois ele conduz somente
até onde os outros chegaram”
Alexander Graham Bell*

AGRADECIMENTOS

Ao Professor Doutor Fernando Lopes, pela disponibilidade, dedicação e apoio prestados na elaboração teórica e prática do Trabalho de Projeto.

Ao Professor Doutor Inácio Fonseca, pela disponibilidade e colaboração na componente prática do projeto.

A todos os Professores do Instituto Superior de Engenharia de Coimbra que contribuíram com o seu vasto conhecimento para a minha formação académica.

Aos meus amigos, Engenheiros Eletrotécnicos, David Gonçalves, Tiago Vasco, Luís Rodrigues e Tiago Duarte, pelo apoio disponibilizado ao longo do meu percurso académico.

Ao Instituto Superior de Engenharia de Coimbra, onde fiz o meu percurso académico.

Aos amigos que estiveram sempre presentes, Márcio Jordão, Geovanne Bulla, Milton Neves, Rafael Guardado, Renato Aparicio, Mauro Nunes, Hugo Sousa, Ângela Pinheiro e Jéssica Vital. Sem vocês Coimbra não teria o mesmo encanto.

À Joana Pedrosa, pelo apoio e compreensão constantes durante toda a elaboração do Trabalho de Projeto.

Por último, mas não em último, aos meus pais, irmão, cunhada e sobrinhos, pelo indubitável apoio que disponibilizaram ao longo de todo o meu percurso académico. Aos meus avós, que, mesmo não estando presentes, me deram forças durante todos os dias.

RESUMO

Este trabalho teve como objetivo principal o estudo da monitorização de áreas ou volumes nos mais diversos ambientes. Um objetivo mais específico foi o desenvolvimento de um sistema de monitorização de segurança e proteção em ambiente industrial utilizando sensores RGB-D.

Procedeu-se ao levantamento da legislação em vigor para a implementação e utilização de sistemas de monitorização de espaços públicos e privados com base em vídeo, bem como das tecnologias e sistemas existentes atualmente no mercado com o mesmo objetivo. Relativamente à legislação a considerar, foi necessário analisar duas perspetivas diferentes. A perspetiva da videovigilância e monitorização de espaços com o objetivo de garantir a segurança de pessoas e bens em relação a comportamentos impróprios e criminais, como vandalismo, roubo e violência, que obedece a regras definidas em legislação própria para a sua implementação e utilização. A outra perspetiva refere-se ao ambiente industrial, onde os sistemas de monitorização têm por finalidade o apoio à produção e à segurança das pessoas e equipamentos, e onde os aspetos mais relevantes a considerar estão associados ao cumprimento de normas relativas aos procedimentos, equipamentos e instalações. Estas duas perspetivas podem conjugar-se se existir captura e gravação de imagens e vídeo, onde seja permitida a identificação dos locais e das pessoas.

Em termos tecnológicos, começou-se por analisar as técnicas e tecnologias associadas à captação de imagem e vídeo 3D, tais como a Triangulação, *Stereo Vision* e a técnica *Time Of Flight*. Considerando a técnica *Time Of Flight*, a mais avançada e com melhores resultados em termos de exatidão e precisão, procedeu-se à pesquisa de *hardware* que satisfizesse as necessidades específicas do projeto. O sensor Kinect V2 da Microsoft satisfaz essas necessidades e, portanto, foi a escolha natural para avançar com o projeto.

Relativamente ao *software* utilizou-se o Matlab como ambiente de desenvolvimento que, em conjunto com o SDK da Microsoft e código C específico de integração do Kinect V2 no Matlab, serviram de base para a evolução e a concretização dos objetivos do projeto. No desenvolvimento da aplicação foram consideradas duas fases, a fase de calibração e a fase de monitorização e deteção. Na fase de calibração foi definido o volume a monitorizar e quantificada a ocupação do volume sem a existência de elementos invasivos. Na fase de monitorização e deteção foi capturada a informação 3D, que permite identificar elementos externos que representam uma intrusão ou evasão no/do volume, que está constantemente a ser monitorizado. Este processo de monitorização foi conseguido através da implementação de um algoritmo de quantificação do volume ocupado, baseado na Triangulação Delaunay em 3D.

Para proceder ao teste em ambiente industrial do sistema desenvolvido, este foi inserido no espaço de operação de um braço robótico a funcionar em modo contínuo, de modo a verificar o seu bom funcionamento e avaliar a sua performance.

Por fim, como objetivo de desenvolvimento futuro, foi proposta a associação integrada de vários sistemas Kinect V2, de forma a mapear completamente todo o volume 3D que se pretende monitorizar, possibilitando total flexibilidade de monitorização e deteção.

Palavras-chave: Dispositivos de Proteção Baseados em Visão; Segurança em Ambiente Industrial; Aquisição de Imagem e Vídeo 3D; Sensores RGB-D, *Time Of Flight*; Kinect V2; Triangulação Delaunay.

ABSTRACT

This work had as general objective the study of the monitoring of areas or volumes in the most diverse environments. A more specific objective was the development of an industrial safety monitoring system using RGB-D sensors.

The legislation in force for the implementation and use of monitoring systems for public and private spaces based on video, as well as technologies and systems currently in the market for this purpose, were surveyed. Regarding the legislation to be considered, it was necessary to analyze two different perspectives. The perspective of video surveillance and space monitoring, in order to ensure the safety of people and property in relation to inappropriate and criminal behavior such as vandalism, robbery and violence, that obeys rules defined in legislation for its implementation and use. Another perspective refers to the industrial environment, where monitoring systems have the purpose of supporting production and the safety of people and equipment, where the most relevant aspects to consider are compliance with rules regarding procedures, equipment and facilities. These two perspectives can be combined, if there is recording of images, where identification of places and people is allowed.

In terms of technology, we began by analyzing techniques and technologies associated with 3D image and video capture, such as Triangulation, Stereo Vision and the Time Of Flight technique. Considering Time Of Flight, the most advanced technique and the best in terms of accuracy and precision results, the next step was to search for hardware that meets the specific needs of the project. The Microsoft Kinect V2 sensor meets these needs and was thus a natural choice to move forward with the project.

Regarding the software, Matlab was used as the development environment, which, together with Microsoft SDK and specific C code for the integration of Kinect V2 in Matlab, served as a basis for the evolution and achievement of the project objectives. For the development of the application, two phases were considered, the calibration phase and the monitoring and detection phase. In the calibration phase, the volume to be monitored was defined and the occupation of the volume without the existence of invasive elements was quantified. In the monitoring and detection phase, 3D information was captured to identify external elements that represent an intrusion or an evasion in/from the volume that is being monitored. The detection process is achieved through the implementation of a Delaunay Triangulation algorithm in 3D, that allows for the quantification of the occupied volume.

In order to test the developed system in an industrial environment, it was inserted in the operating space of a robotic arm working in continuous mode, so as to be able to verify its good functioning and to evaluate its performance.

Finally, for future developments, the integrated association of several Kinect V2 systems is proposed, in order to fully map the entire 3D volume that is intended to be monitored, allowing total flexibility for monitoring and detection.

Keywords: Vision-Based Protective Devices; Safety in the Industrial Environment; 3D Image and Video Acquisition; RGB-D Sensors; Time Of Flight; Kinect V2; Delaunay Triangulation.

ÍNDICE

AGRADECIMENTOS	i
RESUMO	iii
ABSTRACT	v
ÍNDICE DE FIGURAS	ix
SIMBOLOGIA E ABREVIATURAS	xv
1 Introdução	1
1.1 Objetivos	1
1.2 Estrutura do relatório	2
2 Segurança e áreas interditas	3
2.1 Segurança em áreas de acesso restrito	3
2.2 Monitorização em espaços e edifícios usando imagem e vídeo	6
2.3 Aspetos genéricos de segurança na indústria	8
2.3.1 Cultura de segurança	9
2.3.2 Erros e falhas humanas	10
2.4 Legislação sobre sistemas de videovigilância e segurança na indústria	11
2.4.1 Norma EN/IEC61496-1 (Dispositivos de proteção com ação sem contacto)	12
2.4.2 Capacidade de deteção dos dispositivos de proteção optoelectrónicos	17
2.4.3 Distância mínima e tempo de paragem/desaceleração	18
2.4.4 IEC TS 61496-4-2 (VBPD e VBPDPP)	19
2.4.5 Outros requisitos especiais	20
2.5 Tecnologias usadas na definição e restrição de acesso	20
2.5.1 Dispositivos de proteção electro-sensíveis	20
2.5.2 Tecnologias para equipamento de proteção electro-sensível	21
2.6 Aplicação de imagem e vídeo 3D na indústria	25
2.6.1 Sistema SafetyEYE (Pilz)	25
2.7 Outras áreas de aplicação de imagem e vídeo 3D	28
2.7.1 Sistema de monitorização de zona 3D QuadCam	29
3 Aquisição e modelização 3D com sensores RGB-D	31
3.1 Geometria e captura de imagem e vídeo 3D	31
3.2 Técnicas de aquisição de imagem 3D	32
3.2.1 <i>Stereo Vision</i>	32

3.2.2	Triangulação <i>laser</i>	33
3.2.3	<i>Time Of Flight</i>	34
3.3	Geometria e características de sensores RGB-D	36
3.4	Geometria e captura detalhada com o sensor Kinect V2	38
3.4.1	Sensor de Profundidade	38
3.4.2	Transformação de profundidade em coordenadas 3D	43
3.4.3	Alinhamento do sensor de cor e profundidade	46
3.4.4	Aspetos a ter em consideração no funcionamento.....	47
4	Monitorização de áreas e volumes usando o Kinect V2	51
4.1	Especificações do projeto	51
4.2	Fase de calibração do volume interdito	52
4.3	Fase de monitorização	56
4.3.1	Método de Triangulação de Delaunay	56
4.3.2	Técnica de monitorização	58
4.3.3	Aplicação baseada na fase de monitorização	59
4.4	Fluxograma geral do <i>software</i>	65
4.5	Performance temporal do <i>software</i> e <i>hardware</i>	66
4.5.1	Resultados temporais da aquisição de dados	66
4.5.2	Resultados temporais do ciclo de monitorização	68
4.5.3	Performance do computador utilizado.....	70
4.6	Implementação em ambiente industrial.....	71
4.6.1	Resultados da implementação em ambiente industrial.....	72
4.6.2	Condições dos resultados obtidos na implementação em ambiente industrial	75
4.7	Interface com o utilizador	75
4.8	Implementação de alarme.....	77
4.8.1	Método de alarme	77
5	Conclusão e trabalhos futuros	79
5.1	Conclusão	79
5.2	Proposta de desenvolvimento futuro	80
	Referências Bibliográficas	81
	Anexo I Requisitos legais mínimos para os sistemas de videovigilância.	

ÍNDICE DE FIGURAS

Figura 1. Câmera de monitorização de entradas e saídas numa loja [3].	7
Figura 2. Identificação de bagagem abandonada em aeroporto [5].	7
Figura 3. Câmeras de monitorização instaladas nas ruas [6].	8
Figura 4. Triângulo organizacional [10].	9
Figura 5. Tipos de erros humanos, baseado em [15].	11
Figura 6. Opções disponíveis para a escolha de um ESPE [21].	16
Figura 7. Diferenças principais entre AOPDs de Tipo 2 e Tipo 4 [24].	17
Figura 8. Parâmetros para determinar a distância mínima exigida [24].	18
Figura 9. Perigos onde dispositivos de proteção electro-sensível podem ser usados [24].	21
Figura 10. Perigos em que a proteção electro-sensível não pode ser utilizada [24].	21
Figura 11. Proteção de acesso usando um dispositivo de segurança de feixe de luz [24].	22
Figura 12. Proteção de área usando uma cortina de luz de segurança [24].	22
Figura 13. Proteção de área de perigo com um <i>scanner</i> de <i>laser</i> de segurança [24].	23
Figura 14. Modelo de deteção de <i>scanner laser</i> de segurança [24].	24
Figura 15. Estrutura básica de um <i>scanner a laser</i> [24].	24
Figura 16. Proteção com câmara de segurança em célula de robô [24].	25
Figura 17. Sistema de segurança 3D SafetyEYE da Pilz [26].	26
Figura 18. Áreas de proteção do operador, aplicadas a uma máquina [27].	26
Figura 19. Área de trabalho de um robô numa linha de montagem [27].	27
Figura 20. Criação de áreas de proteção através do configurador do SafetyEYE [28].	27
Figura 21. Demonstração das áreas abrangidas pelo SafetyEYE [26].	28
Figura 22. Sistema QuadCam [29].	29
Figura 23. Monitorização de áreas com o sistema QuadCam [29].	29
Figura 24. Possível área de monitorização com o sistema QuadCam [29].	30
Figura 25. Relação dos métodos óticos em 3D, com base em [30].	31
Figura 26. Método de Triangulação <i>Stereo Vision</i> [32].	33
Figura 27. Câmera de <i>Stereo Vision</i> [33].	33
Figura 28. Método de Triangulação <i>laser</i> para aquisição de dados 3D [30].	34
Figura 29. Medição da distância com a tecnologia TOF [38].	35
Figura 30. Câmara de tempo de propagação do sinal (TOF) [41].	36
Figura 31. Localização dos <i>lasers</i> e sensores do Kinect V2 [45].	39

Figura 32. Parâmetros do sensor de profundidade [46].	39
Figura 33. Princípio de funcionamento da tecnologia TOF [47].	40
Figura 34. Relação fase-distância para as três frequências de modulação usadas [47],[49].	41
Figura 35. Cálculo da profundidade através da distância [47].	43
Figura 36. Modelo de câmera <i>pinhole</i> [52].	44
Figura 37. Distorções comuns da lente: barril (esquerda), almofada (direita) [47].	46
Figura 38. Profundidade e cor da câmera "pontos cegos" [47].	47
Figura 39. Valor das medições de distância em função do tempo [57].	48
Figura 40. Desvio das distâncias medidas relativamente às medidas reais [57].	49
Figura 41. Ilustração da conexão do Kinect V2 ao computador.	51
Figura 42. Distância dos objetos à câmera.	52
Figura 43. Método de calibração definido pelo utilizador.	53
Figura 44. Volume interdito definido pelo utilizador.	53
Figura 45. Zona de perigo em torno do espaço interdito.	54
Figura 46. Fluxograma da fase de calibração.	55
Figura 47. Circunferências com Critério de Delaunay [60].	56
Figura 48. Circunferências com Critério de não-Delaunay [60].	56
Figura 49. Triangulação de Delaunay 3D, composta por dois tetraedros [60].	57
Figura 50. Divisão do volume interdito em tetraedros.	58
Figura 51. Fluxograma da fase de monitorização.	59
Figura 52. Funcionamento da fase de monitorização (ciclo em curso).	60
Figura 53. Funcionamento da fase de monitorização (zona de perigo infringida).	60
Figura 54. Funcionamento da fase de monitorização (zona interdita infringida).	61
Figura 55. Funcionamento da fase de monitorização (objeto dentro da zona interdita).	62
Figura 56. Funcionamento da fase de monitorização (objeto retirado da zona interdita).	62
Figura 57. Funcionamento da fase de monitorização (zona dinâmica).	63
Figura 58. Funcionamento da fase de monitorização (intrusão na zona dinâmica).	64
Figura 59. Fluxograma geral do programa.	65
Figura 60. Tempo necessário para que o Kinect V2 responda a uma solicitação de dados.	67
Figura 61. Tempo necessário para a aplicação executar um ciclo de monitorização completo.	69
Figura 62. Robô Motoman usado para a monitorização.	71
Figura 63. Volume interdito definido.	72

Figura 64. Monitorização em curso.....	73
Figura 65. Volume de perigo infringido.....	73
Figura 66. Volume interdito infringido.	74
Figura 67. Relatório gerado quando há intrusão/evasão no/do espaço interdito.....	74
Figura 68. Seleção do modo de funcionamento.	75
Figura 69. Iniciar o modo de monitorização.....	76
Figura 70. Relatório de ocorrências e gravações.....	76
Figura 71. Ilustração da ligação do arduino ao Matlab.	77
Figura 72. Fluxograma do sistema de alarme.....	78
Figura 73. Monitorização utilizando vários sistemas Kinect V2, adaptado de [68].....	80

ÍNDICE DE TABELAS

Tabela 1. Relação dos tipos para PL ou SIL [21].....	13
Tabela 2. Valores limite de probabilidade de falha perigosa por hora para SIL e PL [21].....	13
Tabela 3. Máximo atingível PL e classificação SIL [21].	14
Tabela 4. Características das diferentes marcas de câmeras 3D.....	37

SIMBOLOGIA E ABREVIATURAS

2D – Duas dimensões

3D – Três dimensões

AODP – *Active Opto-electronic Protective Device*

AOPDDR – *Active Opto-electronic Protective Devices Responsive to Difuse Reflection*

CENELEC – Comité Europeu de Normalização Eletrotécnica

CMOS – *Complementary Metal-Oxide-Semiconductor*

CNPD – Comissão Nacional de Proteção de Dados

ESPE – *Electro-Sensitive Protective Equipment*

IEC – *International Electrotechnical Commission*

IP20 – Índice de Proteção 20

IP65 – Índice de Proteção 65

IR – *Infrared*

LED – *Light Emitting Diode*

NUI – *Natural User Interface*

OSSDs – *Output Signal Switching Devices*

PFHd – *Probability of Dangerous Failure per Hour*

PL – *Performance Level*

RGB-D – *Red-Green-Blue – Depth*

SDK – *Software Development Kit*

SIGESP – Sistema Integrado de Gestão de Segurança Privada

SIL – *Safety Integrity Level*

TOF – *Time Of Flight*

USB – *Universal Serial Bus*

VBPDs – *Vision Based Protective Devices*

VBPDPPs – *Vision Based Protective Devices using Passive reference Pattern techniques*

VBPDST – *Vision Based Protective Devices using Stereo vision Techniques*

1 Introdução

Os primeiros sistemas de videovigilância surgiram por volta do ano de 1965 e eram utilizados pela polícia para manter locais públicos vigiados, requerendo observação constante. Em 1970 introduziram-se os sistemas de vídeocassetes, permitindo assim, a gravação em locais como bancos, lojas e bombas de gasolina. Apenas na década de 1990, com o surgimento dos multiplexers digitais, foi possível monitorizar várias câmeras, começando então a ser instaladas em terminais multibanco, com o objetivo de garantir a segurança das transações realizadas pelos clientes.

Com o evoluir da tecnologia e a massificação da era digital, mais propriamente na indústria, os sistemas de videovigilância deixaram de ser meramente sistemas de monitorização passivos e começaram a ser instalados os primeiros sistemas ativos, que interagem com o funcionamento das máquinas, permitindo desta forma evitar inúmeros acidentes.

Apesar da maioria dos sistemas atualmente implementados utilizar a tecnologia 2D, na indústria, nomeadamente da área de controlo de qualidade, já são utilizados inúmeros sistemas de monitorização de imagem, que tiram partido das características que a tecnologia 3D oferece.

Relativamente às técnicas utilizadas para aquisição de dados em 3D, estas têm de ser consideradas consoante a especificidade da aquisição que se pretende obter. Isto deriva do facto de cada técnica utilizar um método próprio que, em determinadas circunstâncias, poderá não satisfazer a funcionalidade para a qual foi escolhida. Serão abordadas as técnicas mais relevantes de aquisição de dados de imagem e vídeo 3D, com uma apresentação mais detalhada da técnica utilizada no desenvolvimento deste projeto, o *Time Of Flight* (TOF).

Por fim, este projeto aborda o que se considera ser a próxima revolução tecnológica industrial, que certamente passará pela implementação de câmeras 3D capazes de fazer uma gestão inteligente dos dados de imagem e vídeo adquiridos, e, desta forma, contribuir para a tendência industrial futura, a designada Indústria 4.0.

1.1 Objetivos

Este projeto teve como principal objetivo a monitorização de áreas ou volumes nos mais diversos ambientes, mais especificamente em ambientes industriais. Esta monitorização foi realizada recorrendo à técnica TOF, utilizando o sensor Kinect V2 da Microsoft.

Deste modo, e de forma a promover a escolha adequada do equipamento a utilizar, foi efetuado o levantamento das tecnologias existentes e, de seguida, a comparação das características das principais câmaras disponíveis.

De modo a proceder ao desenvolvimento do projeto, foi necessário enquadrá-lo nos requisitos teóricos e legais, que são fundamentais para considerar a sua implementação industrial futura. Para isso, abordou-se a legislação sobre implementação de sistemas de videovigilância e analisou-se em pormenor a legislação associada ao ambiente industrial, seja para a monitorização de espaços seja para construção de sistemas e/ou máquinas. Foi necessário recorrer a esta abordagem detalhada em termos de legislação, de forma a poder desenvolver aplicações que combinem a vertente industrial e a vertente de videovigilância.

Para cimentar todo o projeto prático desenvolvido foi elaborado um relatório dos testes efetuados e os respectivos resultados em ambiente real, bem como um sistema de alarme, que permite ao utilizador, através de um alarme luminoso e sonoro, saber com precisão o momento em que houve uma intrusão ou evasão no/do espaço definido previamente.

Pretende-se assim, com recurso a tecnologia avançada mas de custo moderado devido à sua utilização em dispositivos de grande consumo, contribuir para colmatar inúmeras lacunas em termos de segurança existentes atualmente na indústria, bem como contribuir para o objetivo de melhorar a confiança dos trabalhadores e, consecutivamente, a sua produtividade.

1.2 Estrutura do relatório

Este relatório encontra-se dividido em cinco capítulos e um anexo.

O Capítulo 1 (atual) é composto pela introdução, onde é feita uma breve descrição do tema desenvolvido neste projeto, bem como pelos objetivos do trabalho.

O Capítulo 2 foi subdividido em sete subcapítulos. O primeiro subcapítulo faz referência à legislação em vigor que abrange a videovigilância na sua generalidade. O segundo refere a utilização de câmeras de vigilância com monitorização nos mais variados ambientes. O Subcapítulo 2.3 trata os aspetos mais importantes da segurança na indústria, enquanto o Subcapítulo 2.4 introduz a legislação em vigor e os seus aspectos referentes aos sistemas de videovigilância na indústria. Relativamente ao Subcapítulo 2.5, este apresenta as tecnologias usadas na marcação e restrição de acesso em ambiente industrial. O penúltimo subcapítulo, Subcapítulo 2.6, exemplifica um sistema de imagem e vídeo 3D em meio industrial existente no mercado. Por fim, o Subcapítulo 2.7 apresenta uma outra aplicação de imagem e vídeo 3D em ambiente não industrial.

No Capítulo 3 existe uma subdivisão por quatro subcapítulos. O Subcapítulo 3.1 explica a geometria e captura de imagem e vídeo 3D. O segundo subcapítulo faz referência às três principais técnicas de aquisição de imagem 3D, explicando os princípios em que se baseiam. No Subcapítulo 3.3 são abordadas a geometria e as características dos sensores RGB-D, onde é apresentada uma tabela com as principais características de várias câmeras RGB-D. No último subcapítulo (Subcapítulo 3.4) são apresentadas detalhadamente a geometria e a técnica de captura de imagem e vídeo 3D com o sensor Kinect V2, incluindo as relações matemáticas principais que suportam a quantificação dos conceitos de medição, transformação de coordenadas, profundidade, *pinhole*, distorção e alinhamento e, por fim, os aspetos mais práticos a ter em consideração no funcionamento do Kinect V2.

Relativamente ao Capítulo 4, este é subdividido em oito subcapítulos, os quais refletem toda a componente de implementação prática do projeto. Faz-se referência às ligações de *hardware*, bem como ao desenvolvimento do *software*, que, no seu conjunto, constituem um sistema de detecção e protecção de volume em ambiente industrial. Apresentam-se ainda os resultados obtidos a partir de um conjunto de testes efetuados no ambiente de operação de um robô industrial, e, por último, a implementação de um sistema simples de alarme visual e sonoro.

O Capítulo 5 apresenta as conclusões e as propostas de trabalho futuro. É ainda incluído, no final, o Anexo I, que diz respeito aos requisitos legais mínimos para os sistemas de videovigilância.

2 Segurança e áreas interditas

A segurança é a percepção de se estar protegido de riscos, perigos ou perdas. Nesse sentido este capítulo aborda a segurança em espaços de acesso restrito, a legislação associada, assim como as principais tecnologias e respetivas implementações utilizadas em ambiente industrial. Relativamente à legislação a ter em conta, é necessário analisar duas perspetivas diferenciadas. A perspetiva da videovigilância e da monitorização de espaços, onde existem regras definidas em Portaria para a sua implementação e operação. E a perspetiva da indústria, com a finalidade de apoio à produção e segurança em ambiente industrial, onde os aspetos a considerar são mais relevantes em termos do cumprimento de normas reguladoras. Estas duas perspetivas podem conjugar-se nos sistemas de apoio à produção e segurança, se for permitida a gravação de dados com identificação dos operadores. Nesta linha de pensamento, ambas as perspetivas serão abordadas neste capítulo.

2.1 Segurança em áreas de acesso restrito

A segurança tem como prioridade proteger as pessoas das mais variadas ocorrências. Nestes termos foi necessário criar e adaptar a legislação para que esta abranja as entidades no exercício de prestação de segurança. Com a Lei n.º 34/2013, de 16 de maio e a Portaria n.º 273/2013 de 20 de agosto, houve uma importante revisão global do regime jurídico, que regulava o exercício da atividade de segurança privada. Nos seus termos, as entidades consultoras de segurança privada que pretendam elaborar estudos de segurança e projetos de organização de serviços de segurança privada, passam a estar sujeitas a autorização, sucedendo o mesmo com as entidades que procedam à instalação, manutenção ou assistência técnica de material e equipamento de segurança ou de centrais de alarme, sendo obrigatório o seu registo prévio para o exercício da atividade.

Neste contexto, a presente Portaria n.º 273/13 introduz importantes inovações no que toca às condições físicas, materiais e humanas das entidades de segurança privada, adotando um quadro referencial de normas que contribuem para a existência dos necessários padrões mínimos de segurança.

Por fim, e no que se refere a entidades obrigadas a adotar um sistema de segurança ou a imposição de regras de conduta visando a redução de riscos para pessoas e bens e a prevenção da prática de crimes, densificam-se os requisitos dos meios obrigatórios. A presente Portaria regulamenta também aspetos técnicos relacionados com alarmes particulares ou ligados a centrais de receção e monitorização, adequando-os à evolução tecnológica.

De modo a integrar os aspetos mais relevantes desta Portaria no contexto deste trabalho, apresentam-se aqui os artigos mais relevantes:

“Artigo 7.º

Requisitos gerais de segurança das instalações

1 — As instalações operacionais das empresas de segurança privada devem possuir um sistema de segurança físico e eletrónico que compreenda os seguintes requisitos mínimos:

- a) Sistema de videovigilância por câmaras de vídeo para captação e gravação de imagens, com cobertura das áreas de acesso às instalações, que cumpram os requisitos mínimos fixados no Anexo I à presente portaria, do qual faz parte integrante;
- b) Sistema de deteção contra intrusão;
- c) Conexão a uma central de controlo, recetora de sinais de alarme e de videovigilância de funcionamento permanente, com redundância de comunicação e um canal de comunicação que permita transmissão de dados e supervisão permanente de linhas.”

No âmbito deste artigo, são apresentados no Anexo I os requisitos mínimos legais para instalação e funcionamento de sistemas de videovigilância que utilizam câmaras na aquisição e gravação de imagens e vídeo.

“Artigo 31.º

Divulgação e publicidade

1 — A Direção Nacional da PSP assegura na sua página oficial a divulgação das entidades de segurança privada e entidades formadoras autorizadas, por tipo de serviços.

2 — Para efeitos do n.º 5 do artigo 51.º da Lei n.º 34/2013, de 16 de maio, a publicitação de alvarás, licenças, autorizações e respetivos averbamentos, compreende a seguinte informação:

- a) Nome ou designação social e sede;
- b) Número de Identificação de Pessoa Coletiva (NIPC);
- c) Contacto telefónico, fax e email, quando se trate de pessoas coletivas;
- d) Serviços autorizados;
- e) Número, tipo e validade do alvará, licença ou autorização.

3 — Após a emissão de uma autorização de entidade formadora, a Direção Nacional da PSP deve disponibilizar à entidade autorizada o logótipo de entidade certificada, bem como as regras de utilização que esta deve adotar na sua publicidade.”

De forma a proceder à instalação própria de sistemas de videovigilância, é necessária uma autorização da entidade formadora, que posteriormente é divulgada no site oficial da Direção Nacional da PSP.

“Artigo 51.º**Objeto do registo**

1 - O dever de registo previsto no n.º 1 do artigo 31.º da Lei n.º 34/2013, de 16 de maio, relativo aos sistemas de videovigilância por câmaras de vídeo para captação e gravação de imagem é aplicável aos sistemas utilizados pelas entidades titulares de alvará ou licença.

2 — O dever de registo compreende os sistemas de videovigilância por câmaras de vídeo para captação e gravação de imagens de que as entidades de segurança privada titulares de alvará ou licença sejam responsáveis pelo tratamento de dados pessoais.”

O artigo 51º, refere que apenas as entidades titulares de alvará ou licença são responsáveis pelo tratamento de dados pessoais.

“Artigo 52.º**Conteúdo do registo**

1 — O registo dos sistemas de videovigilância por câmaras de vídeo para captação e gravação de imagens é efetuado através do SIGESP, mediante pedido apresentado pela entidade de segurança privada titular de alvará ou licença, ou seu representante.

2 — O pedido de registo deve conter as seguintes informações:

- a) Identificação da autorização da CNPD;
- b) Nome e endereço do responsável pelo tratamento de dados e, se for o caso, do seu representante;
- c) Finalidades do tratamento;
- d) Características do sistema de videovigilância;
- e) Descrição sumária das medidas de segurança físicas e lógicas do tratamento em aplicação dos artigos 14.º e 15.º da Lei n.º 67/98, de 26 de outubro.

3 — Após a submissão do registo a Direção Nacional procede à análise do pedido e emite comprovativo de registo do sistema de videovigilância.”

O pedido de registo por parte da entidade da segurança privada titular de alvará ou licença deve ser efetuado através do SIGESP, e conter as informações requeridas no ponto 2 do artigo 52.º, após uma análise por parte da Direção Nacional é emitido ou não o registo do sistema de videovigilância.

“Artigo 115.º

Sinalização de sistemas de videovigilância

1 — O símbolo identificativo a utilizar na identificação dos locais objeto de vigilância com recurso aos meios previstos no n.º 1 do artigo 31.º da Lei n.º 34/2013, de 16 de maio, constam do anexo VIII à presente portaria, da qual faz parte integrante.

2 — Os requisitos e especificações técnicas da sinalização e as suas dimensões devem cumprir as disposições da norma ISO 3864-1.

3 — O aviso a que se refere o n.º 5 do artigo 31.º da Lei n.º 34/2013, de 16 de maio, deve ser colocado de forma a garantir boas condições de legibilidade das mensagens nele contidas e a acautelar a normal circulação e segurança dos utentes dos espaços.

4 — Os avisos são colocados no perímetro exterior do local ou zona objeto de vigilância com recurso a equipamentos eletrónicos de videovigilância por câmaras de vídeo, e da forma mais conveniente ao seu pronto reconhecimento pelos utentes.

5 — No interior do local ou zona objeto de vigilância devem ser repetidos os avisos de informação.”

O objetivo do artigo 115.º é elucidar acerca da sinalização dos sistemas de videovigilância, tendo em conta o Anexo VIII e a norma ISO 3864-1 [1].

2.2 Monitorização em espaços e edifícios usando imagem e vídeo

Atualmente é cada vez mais comum, em ambientes fechados e também nas ruas das principais cidades, apostar em sistemas de visão inteligentes, que monitorizam um conjunto de parâmetros, sempre com a finalidade da proteção do ser humano. É baseado nestas premissas que se baseiam os exemplos de sistemas de monitorização em centros comerciais, aeroportos e ruas de cidades muito movimentadas.

No setor comercial existe uma forte adesão à tecnologia oferecida pelas câmaras de vigilância que permitem, com base em algoritmos específicos, por exemplo, contar o número de pessoas que passam numa determinada área (Figura 1), elaborando desta forma mapas de acordo com as tendências de ocupação, que se tornam fundamentais em termos de marketing [2].



Figura 1. Câmera de monitorização de entradas e saídas numa loja [3].

No setor aeroportuário existe também um crescente recurso a sistemas de videovigilância inteligentes, que permitem uma deteção automática de comportamentos suspeitos e a identificação de bagagens abandonadas (Figura 2). As câmeras de monitorização podem ajudar a verificar o ocorrido, alertando o operador conforme necessário. Uma função de "pesquisa forense" permite uma recuperação rápida de sequências de vídeo relevantes [4].



Figura 2. Identificação de bagagem abandonada em aeroporto [5].

Nas cidades mais desenvolvidas para fazer face a comportamentos menos próprios, começam a ser instaladas câmeras com *software* que permite reconhecimento facial (Figura 3), mas também servem para deteção de objetos abandonados, bem como monitorização de trânsito.

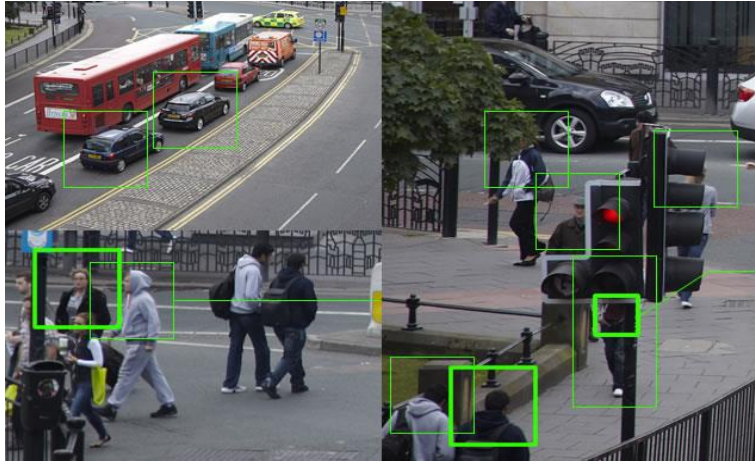


Figura 3. Câmeras de monitorização instaladas nas ruas [6].

Tendo em conta as áreas anteriormente referidas e as empresas que se dedicam à monitorização através de câmeras 2D, é possível enumerar os aspetos mais importantes que esta tecnologia envolve. Assim, consideram-se os seguintes aspetos de maior relevância [7]:

- Identificação e correspondência de veículos;
- Identificação de pessoas;
- Reconhecimento automatizado de matrículas;
- Taxas de transferência;
- Gestão de filas;
- Contagem de pessoas;
- Detecção e classificação de objetos;
- Movimento direcional suspeito;
- Detecção de objetos abandonados;
- Defesa de perímetros.

Tal como vimos nos exemplos referidos anteriormente (lojas, aeroportos e ruas de cidades), na maioria das aplicações, cada um dos itens mencionados surge associado a outros itens, de forma a tornar a aplicação mais abrangente.

2.3 Aspetos genéricos de segurança na indústria

Desde há muito tempo que a segurança na indústria tem vindo a ser um assunto de extrema importância em toda a sociedade, tendo como objetivo principal a aplicação de um conjunto de medidas que zelam pela proteção do trabalhador na sua atividade laboral, assim como a minimização dos acidentes de trabalho.

Para que o desenvolvimento da segurança no trabalho, em meio empresarial, seja eficaz, é importante monitorizar um conjunto de processos, de onde se destacam, o sistema, a atividade,

as instalações e o meio ambiente. Este controlo é feito através de um sistema organizacional, que não é simples de conceber por si só, mas que considera toda a envolvente.

Assim, com a aplicação de um sistema de segurança no trabalho, tenta-se minimizar todos os possíveis riscos e, apesar de não ser uma tarefa fácil, a sua supressão. A integridade da vida humana e a proteção do meio ambiente envolvente têm de estar obrigatoriamente num elevado patamar de proteção. Portanto, deve existir um processo contínuo de supressão ou mitigação dos riscos.

Apesar disso, sabe-se que para que os métodos de segurança sejam eficazes, diminuindo assim a probabilidade de danos para o ser humano, património e meio ambiente, é necessário que as organizações permitam e que a sociedade possua uma mente aberta para aprender e adquirir novos conhecimentos de forma a alargar horizontes e evoluir.

2.3.1 Cultura de segurança

O conceito de cultura de segurança surgiu pela primeira vez na literatura em 1986, após o desastre de Chernobyl, e seguidamente com os acidentes de Piper Alpha e Kings Cross [8]. Após estes acontecimentos, colocou-se em causa a eficácia das organizações no que diz respeito à segurança e saúde no local de trabalho. Assim, de modo a proceder à prevenção dos potenciais acidentes de trabalho, doenças e outras situações consideradas catastróficas, provenientes não só de erros técnicos ou individuais, mas também de fatores como atitudes, perceções e comportamentos dos gestores e dos funcionários, estabeleceu-se o conceito de cultura de segurança, com a finalidade de desenvolver a proteção da saúde e segurança nas organizações [8], [9].

Relativamente à cultura de segurança deduz-se que deve existir um ciclo de interligação, sendo este constituído por um triângulo organizacional em que se insere a estrutura e os métodos da organização, que influenciam a cultura e são refletidos nos comportamentos e desempenho (Figura 4) [10], [11].

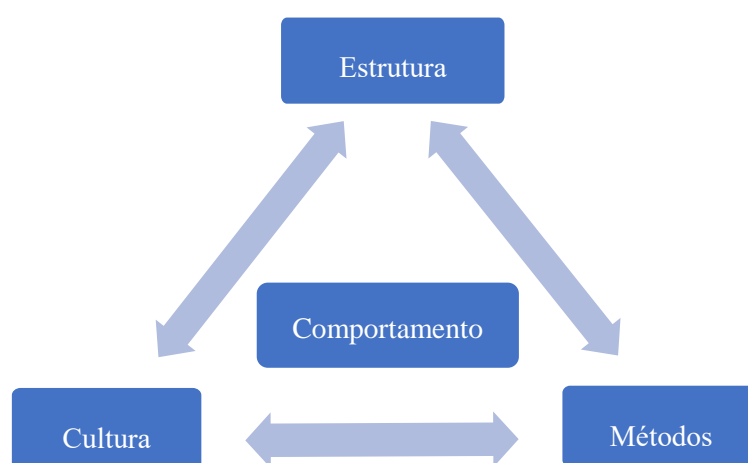


Figura 4. Triângulo organizacional [10].

A cultura de segurança de uma organização deve ser cuidadosamente estudada e constantemente melhorada, para que desta forma possa influenciar o cumprimento das regras

de segurança, promover os atos seguros e desenvolver uma maior motivação que se reflita na segurança, e assim permitir que sejam alteradas as influências negativas, falhas de segurança, incidentes e acidentes no local de trabalho [10], [12], [13].

2.3.2 Erros e falhas humanas

A questão do erro humano é complexa e não pode ser apenas definida como uma desatenção e/ou negligência do trabalhador, pois o comportamento do ser humano num determinado sistema deve ser analisado de uma forma mais detalhada e global.

Um sistema pode ser composto pelo homem e pelo equipamento, mas no desenvolvimento de um processo o homem pode não apresentar um papel direto, o que leva a que não seja ele a provocar o erro.

Sempre foram implementadas estratégias para que a frequência de acidentes diminua, através da mudança dos comportamentos dos trabalhadores, com a finalidade de redução do erro humano.

No mundo da produção, as perturbações, falhas técnicas ou os acidentes remetem, a maior parte das vezes, para uma “falha” ou “erro” humano. No entanto, quando o erro é cometido numa situação em que o homem se encontra para além do seu funcionamento cognitivo não é correto culpá-lo isoladamente por esse facto irremediável. O valor pedagógico do erro é irrefutável, pois ele é parte integrante do processo de aprendizagem, mas claro que esse erro não pode sair dos limites de vigilância do seu controlo, limites que o tornam inofensivo [14].

Deste modo, as falhas humanas (Figura 5) são categorizadas consoante os seus padrões, provenientes de inúmeros fatores que influenciam o desempenho, isto é, carga de trabalho, pressão de tempo, distrações, competência, etc.

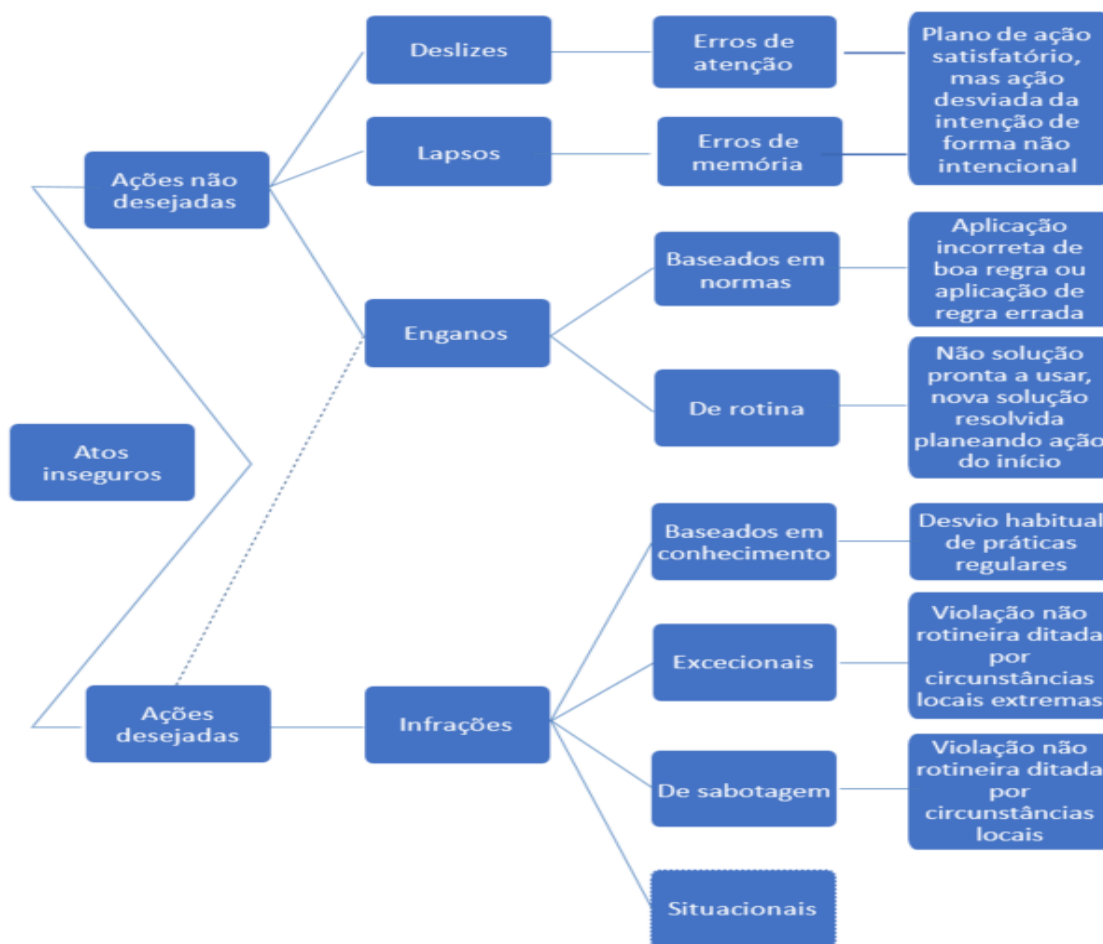


Figura 5. Tipos de erros humanos, baseado em [15].

Considerando o diagrama da Figura 5, pode-se definir o erro humano como resultado da incapacidade de tratamento, ao nível cognitivo, da informação necessária de modo a não errar. Essa situação pode ser provocada pela falta de informação disponível, excesso de informação, ou por fatores externos ao processo. Assim, é importante gerir os erros humanos previsíveis e adotar técnicas de segurança para a redução dos mesmos [15].

2.4 Legislação sobre sistemas de videovigilância e segurança na indústria

A legislação sobre sistemas de videovigilância e segurança na indústria assenta no artigo 20.º do Código de Trabalho Português [16]. Sabe-se que a utilização deste tipo de sistemas deve ter única e exclusivamente a finalidade de garantir a segurança dos trabalhadores e nunca o controlo ou vigilância, o qual constitui uma contraordenação muito grave.

“Artigo 20.º do Código do trabalho

Meios de vigilância à distância:

1 – O empregador não pode utilizar meios de vigilância à distância no local de trabalho, mediante o emprego de equipamento tecnológico, com a finalidade de controlar o desempenho profissional do trabalhador.

2 – A utilização de equipamento referido no número anterior é lícita sempre que tenha por finalidade a proteção e segurança de pessoas e bens ou quando particulares exigências inerentes à natureza da atividade o justifiquem.

3 – Nos casos previstos no número anterior, o empregador informa o trabalhador sobre a existência e finalidade dos meios de vigilância utilizados, devendo nomeadamente afixar nos locais sujeitos os seguintes dizeres, consoante os casos: «Este local encontra-se sob vigilância de um circuito fechado de televisão» ou «Este local encontra-se sob vigilância de um circuito fechado de televisão, procedendo-se à gravação de imagem e som», seguido de símbolo identificativo.

4 – Constitui contraordenação muito grave a violação do disposto no n.º 1 e constitui contraordenação leve a violação do disposto no n.º 3.”

Em suma, qualquer sistema de videovigilância que seja instalado em ambiente industrial, deve obedecer aos requisitos mencionados no artigo 20.º do Código de Trabalho Português.

2.4.1 Norma EN/IEC61496-1 (Dispositivos de proteção com ação sem contacto)

A Comissão Eletrotécnica Internacional (IEC), é uma organização internacional que publica princípios internacionais competentes para as normas nacionais correspondentes.

A IEC 61496-1 [17] especifica os requisitos gerais para a conceção, construção e teste de dispositivos de proteção electro-sensíveis. Esta versão foi publicada em 5 de abril de 2012 e substituiu a IEC 61496-1: 2004-02 (2.ª edição) e a sua emenda AMD1: 2008-01. Na UE, o CENELEC publicou a EN 61496-1: 2013-11 em paralelo com esta. A Comissão Europeia publicou esta versão pela primeira vez no Jornal Oficial da UE em abril de 2014, nos termos da Diretiva Máquinas 2006/42/CE [18]. Em 10 de maio de 2015 foi definido como o fim do período de transição para a presunção de conformidade do documento antecessor EN 61496-1: 2004-05 em conjunto com a alteração EN 61496-1 / A1: 2008-08.

Uma vez determinado o nível de segurança requerido como parte da avaliação de risco, ou *Safety Integrity Level* (SIL), e de acordo com o método numérico da IEC 62061(SIL) [19], ou o gráfico de risco ou *Performance Level* (PL) da norma ISO 13849 [20], surge a questão de como implementar as medidas técnicas de proteção com dispositivos de proteção electro-sensíveis.

A capacidade de deteção descreve a capacidade de detetar de forma confiável objetos de um determinado tamanho, resultando assim na comutação das saídas para o dispositivo de proteção optoelectrónico. É um dos principais parâmetros para a definição de utilização, por exemplo, para detetar dedos, mãos ou pessoas, e para definir a distância mínima ao ponto perigoso. O nível de segurança, com base na classificação do tipo, também leva em conta o grau de confiabilidade da capacidade de deteção.

Os requisitos de fontes óticas de interferência como a luz solar, tipos de lâmpadas diferentes, superfícies refletoras, desalinhamento durante o funcionamento normal ou as propriedades refletoras de vários objetos em *scanners a laser* de segurança desempenham um papel importante na detecção. Outros exemplos de capacidades sistemáticas que também têm de ser considerados englobam o comportamento do dispositivo no caso de uma falha de compatibilidade eletromagnética.

A versão atual da norma IEC 61496-1 [17] e a versão revista da EN 61496-1 [17] definem claramente a relação do tipo com a PL ou a SIL no contexto da proteção do operador. Um dispositivo de proteção electro-sensível ou *Electro-Sensitive Protective Equipment* (ESPE) deve satisfazer ambos os requisitos para o tipo em questão e os do PL e/ou do SIL atribuídos (ver Tabela 1).

Tabela 1. Relação dos tipos para PL ou SIL de acordo com IEC 61496-1 e EN 61496-1 [21].

	Tipo 1	Tipo 2	Tipo 3	Tipo 4
Desempenho de segurança de acordo com IEC 62061 e / ou ISO 13849-1	-	SIL 1 e SIL class 1 e/ou PL c	SIL 2 e SIL class 2 e/ou PL d	SIL 3 e SIL class 3 e/ou PL e

Mesmo que um ESPE de Tipo 2 seja atribuído a SIL 1 ou PL c devido a capacidades sistemáticas, a taxa de falha ou *Probability of Dangerous Failure per Hour* (PFHd) deste subsistema deve ser inferior aos valores-limite representativos (ver Tabela 2). Isso ocorre porque uma cadeia de segurança geralmente consiste em vários subsistemas. O valor PFHd resultante é constituído pela soma dos valores PFHd individuais.

Tabela 2. Valores limite de probabilidade de falha perigosa por hora para SIL e PL [21].

Nível de performance (EN 13849-1)	Probabilidade de falha perigosa por hora (1/h)	Nível de integridade de segurança de acordo com EN IEC 62061
PL b	$3 \times 10^{-6} \leq PFH_D < 10^{-5}$	SIL 1
PL c	$10^{-6} \leq PFH_D < 3 \times 10^{-6}$	SIL 1
PL d	$10^{-7} \leq PFH_D < 10^{-6}$	SIL 2
PL e	$10^{-8} \leq PFH_D < 10^{-7}$	SIL 3

Na prática, a IEC 62061 [19] e a ISO 13849 [20] são frequentemente especificadas para dispositivos de proteção electro-sensíveis, além da IEC 61496 e da EN 61496-1 [17]. Estas duas

normas aplicam-se em segurança de maquinaria e referem-se à IEC 61508 [22], esta mais fundamental em vários pontos. Se um SIL é especificado de acordo com a IEC 61508, isto dá ao utilizador a oportunidade de aplicar um ESPE fora do setor típico da engenharia mecânica.

Se for necessário SIL, SIL *class* 1 ou PLr c, deve ser utilizado um dispositivo de proteção optoelectrónico que corresponda ao Tipo 2 ou superior da norma IEC 61496 [17] ou EN 61496-1 [17] (ver Tabela 3). No caso de uma cortina de luz Tipo 2, por exemplo, um Tipo 4 seria também permitido neste tipo de aplicação, uma vez que as suas capacidades sistemáticas são mais elevadas. O Tipo 3 não é definido para cortinas de luz e não é descrito na seção correspondente da norma IEC 61496-2 [17] ou EN 61496-1 [17]. Exemplos de sensores Tipo 3 são *scanners* de *laser* de segurança ou sistemas de câmeras de segurança. Inversamente, uma cortina de luz Tipo 2 não seria permitida em aplicações que requerem SIL 2, SIL *class* 2 ou PLr d de acordo com o nível de segurança a ser atingido.

Tabela 3. Máximo PL atingível e classificação SIL para cada tipo de acordo com IEC 61496-1 e EN61496-1 [21].

	Tipo 1	Tipo 2	Tipo 3	Tipo 4
Para a função de segurança que envolve um ESPE, o PL ou SIL máximo que o ESPE pode conseguir é o seguinte:	-	SIL 1 e/ou PLr c	SIL 2 e/ou PLr d	SIL 3 e/ou PLr e

A China, o Japão, os EUA e a Austrália geralmente implementam normas IEC sob a forma de padrões nacionais. No passado, os períodos de transição levaram vários anos. As normas IEC podem constituir a base para aprovações, bem como avaliações de responsabilidade a nível nacional. A aceitação nos centros regionais de homologação é efetuada quer diretamente com base na Edição 3 da IEC 61496 [17], quer através das normas nacionais que foram implementadas a partir desta. Isto pode implicar modificações a nível nacional em cada caso.

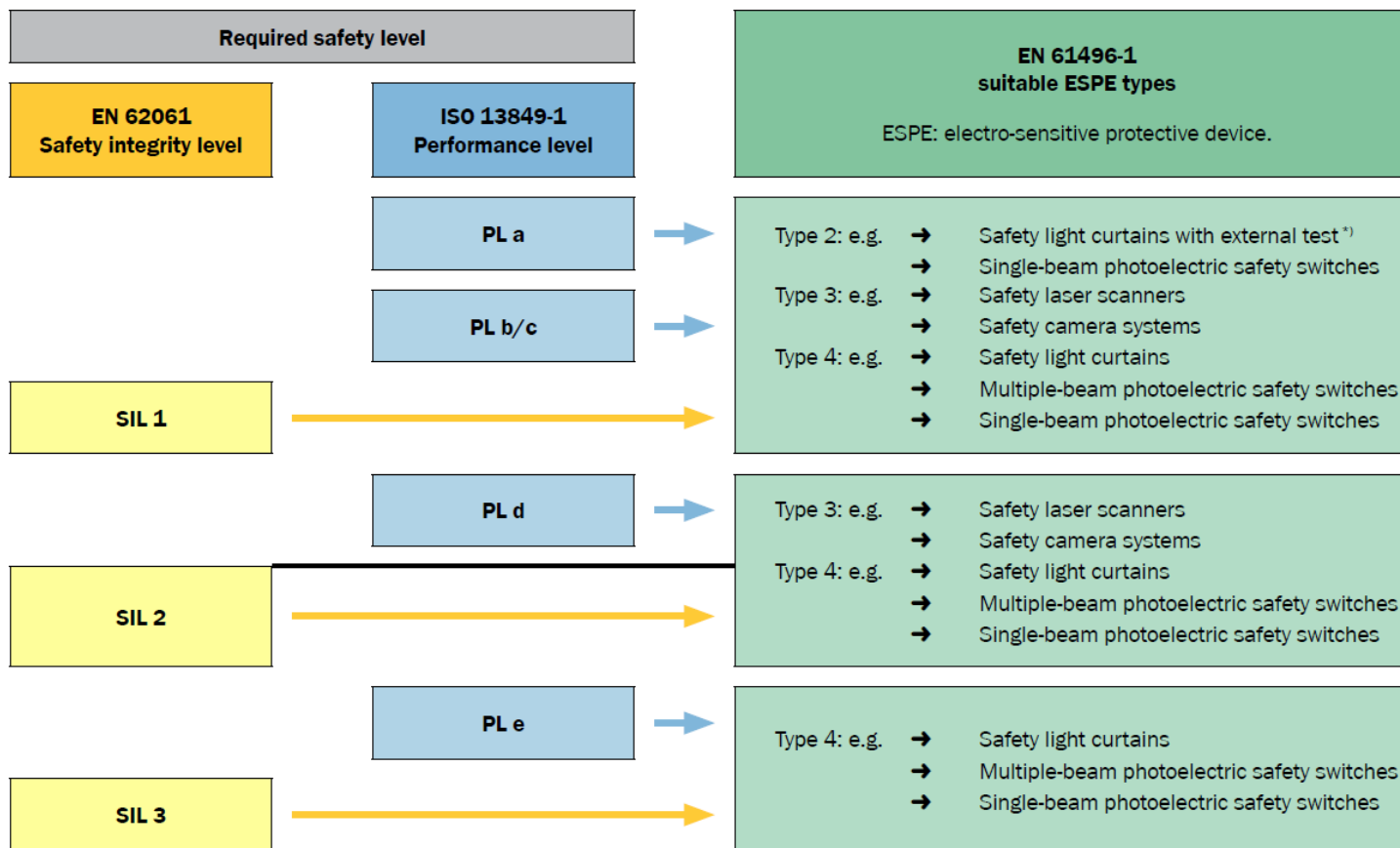
Na Europa, a transição da EN 954 [23] para a EN ISO 13849 [20] para a proteção do operador no domínio da segurança das máquinas conduziu inicialmente a uma discussão sobre como seleccionar uma medida técnica de proteção adequada. Uma vez que, como esperado, a EN 61496-1: 2013-11 [17] segue o modelo que a norma IEC atual definiu, existindo uma clara correlação entre o nível de segurança requerido e o tipo ESPE correspondente.

A aplicação de uma norma como algo que permite presumir conformidade, permanece voluntária. Os fabricantes podem decidir por si se desejam utilizar normas. Contudo, se optarem por não aplicar esta norma, devem provar que as máquinas cumprem os requisitos essenciais da Directiva Máquinas 2006/42/CE [18] através da aplicação de outras especificações.

Além disso, são obrigados a considerar todos os desenvolvimentos significativos no "estado da técnica", o que poderia significar que já não é possível presumir que as máquinas que já possuem um certificado de exame CE de determinado tipo continuam a cumprir determinados requisitos básicos de saúde e segurança. Isso deve ser decidido sempre que a base de uma norma for alterada.

A comprovação de capacidades sistemáticas suficientes de um ESPE numa máquina sem utilizar a norma EN 61496-1 [17] é complexa e envolve custos elevados para o fabricante da máquina. Isto também se aplica às extensões IEC 61496-2 [16] (aplicável a interruptores de segurança fotoelétricos de feixe único, cortinas de luz de segurança) e IEC 61496-3 [17] (aplicável a *scanners a laser* de segurança).

A Figura 6 representa uma versão generalizada das opções disponíveis para a escolha de um ESPE, de acordo com o nível de minimização de risco exigido e com a versão da EN 61496-1 [17].



^{*)} See technical data for the required external tests and their requirement rates.
 Figure 1: Achievable reliability of safety functions with opto-electronic protective devices

Figura 6. Opções disponíveis para a escolha de um ESPE [21].

Portanto, através da Figura 6, podemos constatar que o sistema de câmeras de segurança enquadra-se no Tipo 3 da ESPE (EN61496-1 [17]), com uma probabilidade de falha perigosa por hora enquadrada nos seguintes termos: $10^{-7} \leq PFH_D < 10^{-6}$.

2.4.2 Capacidade de detecção dos dispositivos de proteção optoelectrónicos

A capacidade de detecção é definida como o limite do parâmetro do sensor que faz com que o dispositivo de proteção electro-sensível (ESPE) seja acionado. Na prática, este é o tamanho do menor objeto detetado pelo ESPE dentro da área monitorizada definida (campo de proteção). A capacidade de detecção é especificada pelo fabricante. Em geral, a capacidade de detecção é determinada pela soma da separação do feixe e do diâmetro efetivo do feixe. Isto assegura que um objeto de determinado tamanho interrompe sempre um feixe de luz e é, portanto, detetado independentemente da sua posição no campo de proteção. Para os *scanners a laser* de segurança, a capacidade de detecção é independente da distância ao objeto, do ângulo entre os feixes individuais de luz (pulsos) e a forma e tamanho do feixe transmitido.

A fiabilidade da capacidade de detecção é determinada pela classificação do tipo na norma série EN 61496 [17]. Para o *Active Opto-electronic Protective Device Responsive to Difuse Reflection* (AOPDDR) é definido o Tipo 3. Para o *Active Opto-electronic Protective Device* (AODP) são definidos Tipo 2 e Tipo 4 (Figura 7). Requisitos relativos às fontes óticas de interferência como a luz solar, tipos de lâmpadas diferentes, superfícies refletoras, desalinhamento durante o funcionamento normal e a reflexão difusa dos *scanners laser* de segurança desempenham um papel importante na detecção.

	Type 2	Type 4	Advantage Type 4
Functional safety	Between the test intervals, the protective function may be lost during a failure.	The protective function is retained even during several failures.	Higher risk reduction
EMC (electromagnetic compatibility)	Basic requirements	Increased requirements	Higher reliability of the detection capability
Maximum field of view of the optics	10°	5°	
Minimum distance a to reflective surfaces over a distance D of < 3 m	262 mm	131 mm	Higher system availability in difficult ambient conditions.
Minimum distance a to reflective surfaces over a distance D of > 3 m			
	$= \text{distance} \times \tan(10^\circ/2)$	$= \text{distance} \times \tan(5^\circ/2)$	
Several senders of the same design in a system (workplace)	No special requirements (Beam coding is recommended)	No effect; however, if affected, OSSDs switch off	

Figura 7. Diferenças principais entre AOPDs de Tipo 2 e Tipo 4 [24].

2.4.3 Distância mínima e tempo de paragem/desaceleração

Existem diversos fatores que influenciam a proteção confiável de um ESPE. No nosso contexto justifica-se apenas a integração da distância mínima e do tempo de paragem/desaceleração.

Existe sempre um tempo de paragem/desaceleração após o sinal ser dado para parar as funções perigosas da máquina. O tempo de todo o sistema (toda a cadeia de controlo) está contido neste chamado tempo total de paragem. Este tempo determina a distância mínima necessária do dispositivo de proteção à área de perigo. A distância mínima necessária é calculada de acordo com a norma EN ISO 13855 [25].

A consideração da distância mínima aplica-se a ESPEs com campos de proteção bidimensionais, por exemplo, cortinas de luz (AOPD), *scanners a laser* (AOPDDR) ou sistemas de câmeras bidimensionais.

A fórmula geral para calcular a distância mínima (distância de segurança) é:

$$S = (K \times T) + C \quad (2.1)$$

Onde:

- S é a distância mínima em milímetros, medida no próximo ponto perigoso até ao ponto de deteção e/ou linha de deteção ou plano de deteção do dispositivo de proteção;
- K é um parâmetro em milímetros por segundo, derivado dos dados para as velocidades de aproximação do corpo ou partes do corpo;
- T é o tempo total de paragem do sistema;
- C é uma distância adicional em milímetros.

A distância C depende da capacidade de deteção do sistema (Figura 8). É calculada com base na interrupção da cortina em ângulo reto e depende da altura do campo de proteção acima do nível de referência.

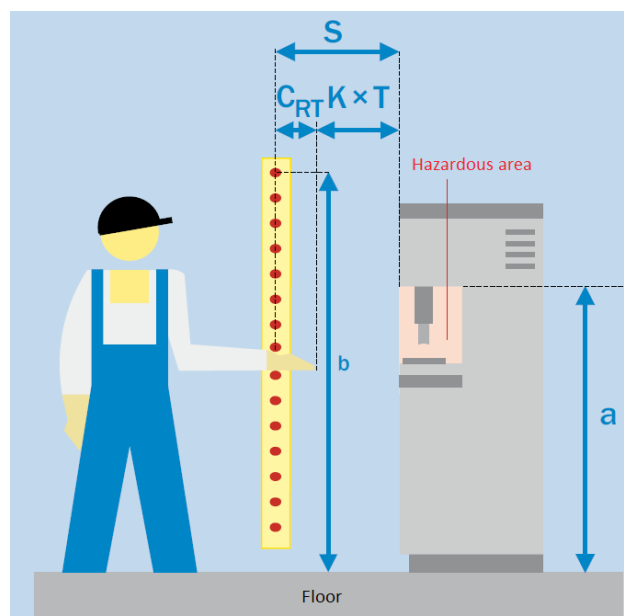


Figura 8. Parâmetros para determinar a distância mínima exigida [24].

Para o tempo de paragem geral T , devem ser considerados os seguintes parâmetros:

- Tempo de paragem da máquina;
- Tempo de resposta do comando de segurança;
- Tempo de resposta do dispositivo de proteção (ESPE);
- Adições de acordo com a capacidade de deteção do ESPE, altura do campo de proteção e/ou tipo de aproximação.

Com a conjugação destes parâmetros pode-se calcular o tempo de paragem/desaceleração.

2.4.4 IEC TS 61496-4-2 (VBPD e VBPDPP)

Esta secção da IEC 61496 [17] especifica os requisitos para a conceção, construção e ensaio de equipamentos de proteção electro-sensíveis (ESPE), concebidos especificamente para detetar pessoas como parte de um sistema de segurança, utilizando dispositivos de proteção baseados na visão ou *Vision Based Protective Devices* (VBPDs) e utilizando padrões passivos de referência ou *Vision Based Protective Devices using Passive reference Pattern techniques* (VBPDPP) para a função de deteção. É dada especial atenção às características que garantem a obtenção de um desempenho adequado, relacionado com a segurança. Um ESPE pode incluir funções opcionais de segurança, cujos requisitos são indicados no Anexo A da IEC-61496-1: 2012 [17].

Ainda assim, esta parte da IEC 61496 [17] não especifica as dimensões ou configurações da zona de deteção e a sua disposição em relação a peças perigosas para qualquer aplicação particular, nem o que constitui um estado perigoso de qualquer máquina. É restrito ao funcionamento do ESPE e como este interage com a máquina.

Um VBPDPP é definido como consistindo em um único dispositivo de deteção de imagem, observando um padrão de referência passivo como fundo e onde o princípio de deteção é baseado no bloqueio ou impedimento parcial da vista do padrão. Informações sobre a espessura, forma, características de superfície ou localização do objeto não é necessária para a deteção. Para dispositivos de deteção de imagens múltiplas, técnicas adicionais, requisitos e procedimentos de teste podem ser necessários:

- Esta parte da IEC 61496 [17] está limitada a ESPEs automáticos baseados em visão que não requerem intervenção humana para deteção;
- É limitado a ESPEs automáticos baseados em visão que detetam objetos que entram ou estão presentes em uma(s) zona(s) de deteção;
- É limitado a ESPEs usando a técnica de iluminação ativa;
- Excluem-se desta especificação técnica os VBPDPPs que utilizam radiação em comprimentos de onda fora do intervalo de 400 nm a 1500 nm;
- Este documento não aborda os aspetos necessários para a classificação ou diferenciação complexa do objeto detetado.

Por fim, esta secção da IEC 61496 [17] é relevante para VBPDPPs com uma capacidade de deteção declarada de até 200 mm.

2.4.5 Outros requisitos especiais

Existe ainda o IEC TS 61496-4-3 [17] - Requisitos especiais para equipamentos que utilizam dispositivos de visão (VBPD) - Requisitos adicionais ao utilizar técnicas de *Stereo Vision* ou *Vision Based Protective Devices using Stereo vision Techniques* (VBPDEST), que tal como o nome indica, faz referência a dispositivos que utilizem a técnica de aquisição *Stereo*, e com a qual se podem obter imagens em 3D.

É de realçar o facto de, até à data, o autor não ter conhecimento da existência de requisitos adicionais para utilização de técnicas de visão TOF.

2.5 Tecnologias usadas na definição e restrição de acesso

As medidas e os produtos para a implementação dos requisitos de segurança das máquinas tornaram-se mais diversificados ao longo dos anos. O objetivo é sempre melhorar a integração da segurança funcional em máquinas e sistemas de salvaguarda. Existem atualmente várias tecnologias para a implementação de medidas de proteção, tal como será tratado de seguida.

2.5.1 Dispositivos de proteção electro-sensíveis

Com dispositivos de proteção electro-sensíveis, ao contrário das barreiras físicas, a proteção não se baseia na separação física de pessoas em risco do próprio risco. A proteção é conseguida através da separação temporal. Enquanto um indivíduo permanecer numa área definida, nenhuma função de máquina perigosa será iniciada e essas funções serão interrompidas, caso se encontrem em andamento. É necessário algum tempo, o designado tempo de paragem/desaceleração, para parar estas funções. O ESPE deve detetar a aproximação de uma pessoa à área de perigo de forma oportuna e, dependendo da aplicação, a presença da pessoa na área de perigo. Os requisitos de segurança para ESPEs, independentemente da sua tecnologia ou princípio de funcionamento, estão estabelecidos na Norma Internacional EN 61496-1 [17].

Se um operador acede regularmente a uma máquina e, portanto, está exposto a um perigo (exemplo da Figura 9), o uso de um ESPE em vez de guardas físicas/mecânicas (coberturas, cercas de segurança, etc.) é vantajoso devido a:

- Tempo de acesso reduzido (o operador não precisa de esperar que o dispositivo de proteção se abra);
- Aumento da produtividade (economia de tempo ao carregar a máquina);
- Melhor ergonomia no local de trabalho (o operador não tem de operar uma guarda física).

Além disso, não só os operadores, mas também outras pessoas estão protegidas.



Figura 9. Perigos onde dispositivos de proteção electro-sensível podem ser usados [24].

Uma vez que os dispositivos de proteção electro-sensíveis não fornecem qualquer barreira física, estes não são capazes de proteger as pessoas contra as emissões, tais como peças de máquinas ejetadas, peças de trabalho ou aparas de metal, radiação ionizante, calor (radiação térmica), ruído, refrigerantes pulverizados, lubrificantes, etc. (Figura 10). O uso de um ESPE também não é possível em máquinas com tempos de paragem/desaceleração longos, que requerem distâncias mínimas irrealizáveis. Nesses casos, devem ser utilizadas barreiras físicas.



Figura 10. Perigos em que a proteção electro-sensível não pode ser utilizada [24].

2.5.2 Tecnologias para equipamento de proteção electro-sensível

Dispositivos de proteção electro-sensíveis podem implementar a deteção de pessoas através de vários princípios: ótico, capacitivo, ultrassónico, micro-ondas e deteção passiva por infravermelho. Devido à precisão reduzida, os sistemas capacitivos e ultrassónicos têm-se mostrado inadequados. Na prática, os dispositivos de proteção optoelectrónicos têm-se mostrado fiáveis ao longo de vários anos, sendo os mais utilizados na indústria, e são descritos como cortinas de luz, *scanners/lasers* e sistemas à base de câmeras, sendo estes descritos com mais detalhe seguidamente.

- **Cortinas de luz de segurança e interruptores fotoelétricos (AOPDs)**

Os *Active Opto-electronic Protective Devices* (AOPDs) são dispositivos de proteção que usam elementos emissores e recetores optoelectrónicos para detetar pessoas numa área bidimensional definida. Uma série de feixes de luz paralelos (habitualmente infravermelhos) transmitidos do emissor para o recetor formam um campo de proteção que salvaguarda a área perigosa. A deteção ocorre quando um objeto opaco interrompe completamente um ou mais feixes. O recetor sinaliza a interrupção do feixe por uma mudança de sinal (estado OFF) para os seus dispositivos de comutação de sinal de saída ou *Output Signal Switching Devices* (OSSDs). Os sinais dos OSSDs são usados para parar as funções perigosas da máquina. A norma internacional IEC 61496-2 [17] inclui os requisitos de segurança para AOPDs.

Os AOPD típicos incluem interruptores de segurança fotoelétricos de feixe único, vários dispositivos de segurança de feixe de luz e cortinas de luz de segurança.

Os dispositivos de segurança de feixe de luz múltiplos são os AOPDs com uma capacidade de detecção de mais de 40 mm. Estes são usados para proteger o acesso a áreas perigosas (Figura 11). Os AOPD com uma capacidade de detecção igual ou inferior a 40 mm são chamados de grelhas de luz de segurança ou cortinas de luz de segurança e são utilizados diretamente para proteger pontos perigosos (Figura 12).

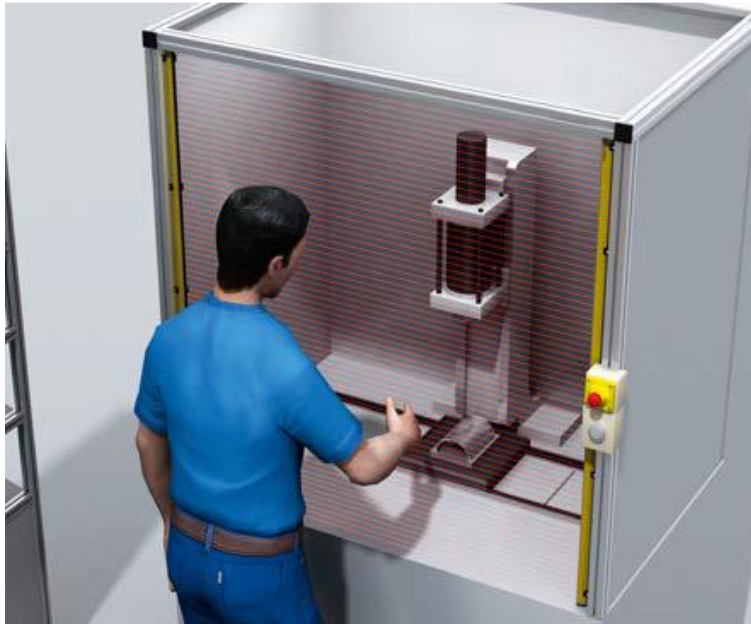


Figura 11. Proteção de acesso usando um dispositivo de segurança de feixe de luz [24].

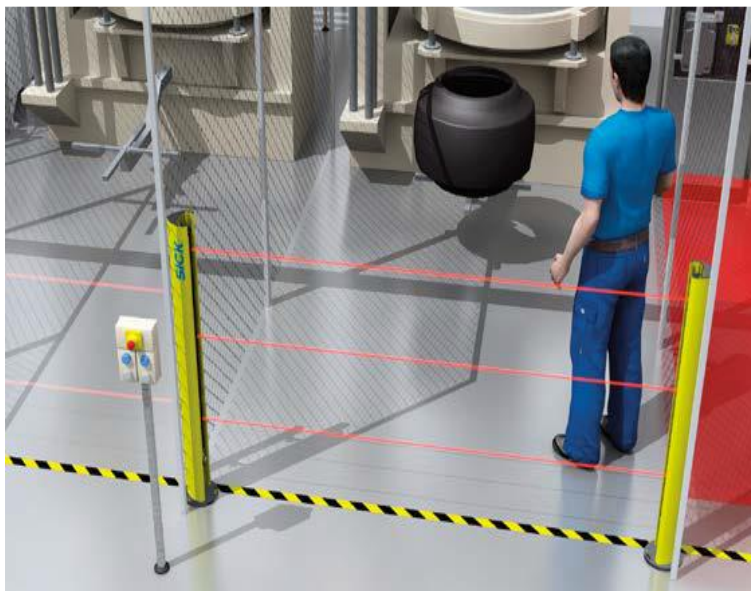


Figura 12. Proteção de área usando uma cortina de luz de segurança [24].

- **Scanners a laser de segurança (AOPDDRs)**

Os *Active Opto-electronic Protective Devices Responsive to Difuse Reflection (AOPDDRs)* são dispositivos de proteção que usam elementos de transmissão e recepção optoelectrónicos para detetar a reflexão da radiação ótica gerada pelo dispositivo de proteção. A reflexão é gerada por um objeto numa área bidimensional definida. A deteção é sinalizada por uma mudança de sinal (estado OFF) para os seus dispositivos de comutação de sinal de saída (OSSDs). Estes sinais dos OSSDs são usados para parar as funções perigosas da máquina.

Os *scanners a laser* de segurança são usados principalmente para proteção de áreas perigosas estacionárias e móveis, como está representado na Figura 13.

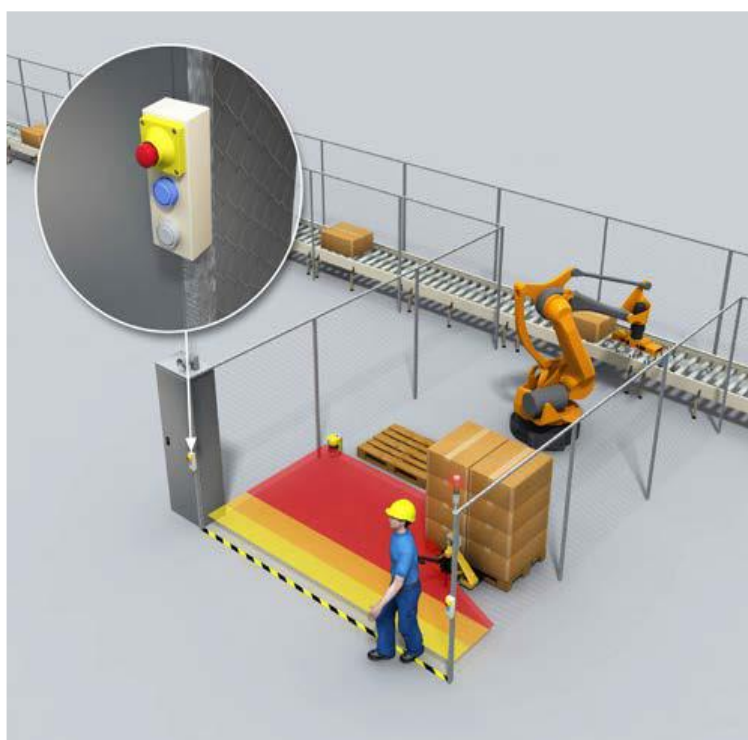


Figura 13. Proteção de área de perigo com um scanner de laser de segurança [24].

O *scanner de laser* de segurança é um sensor ótico, que varre o ambiente com feixes de luz infravermelha em duas dimensões e monitoriza uma área perigosa perto de uma máquina ou veículo. Funciona segundo o princípio da medição do TOF (Figuras 14 e 15). O *scanner* transmite pulsos de luz muito curtos (S), enquanto um "cronómetro eletrónico" é executado simultaneamente. Se a luz atinge um objeto, esta é refletida e recebida pelo *scanner* (R). O *scanner* calcula a distância até ao objeto com base na diferença de tempo entre o emissor e o recetor (Δt). Um espelho girando uniformemente (M) no *scanner* desvia os pulsos de luz de modo a que uma seção de um círculo seja coberta. O *scanner* determina a posição exata do objeto a partir da distância medida e o ângulo de rotação do espelho. O utilizador pode programar a área em que a deteção de objetos dispara o ESPE (campo de proteção). Dispositivos de última geração permitem a monitorização simultânea de várias áreas ou a troca dessas áreas durante o funcionamento. Por exemplo, pode ser utilizado para ajustar a área monitorizada à velocidade do veículo ou uma resposta gradual (campo de aviso - campo de proteção) para evitar interrupções desnecessárias nas operações.

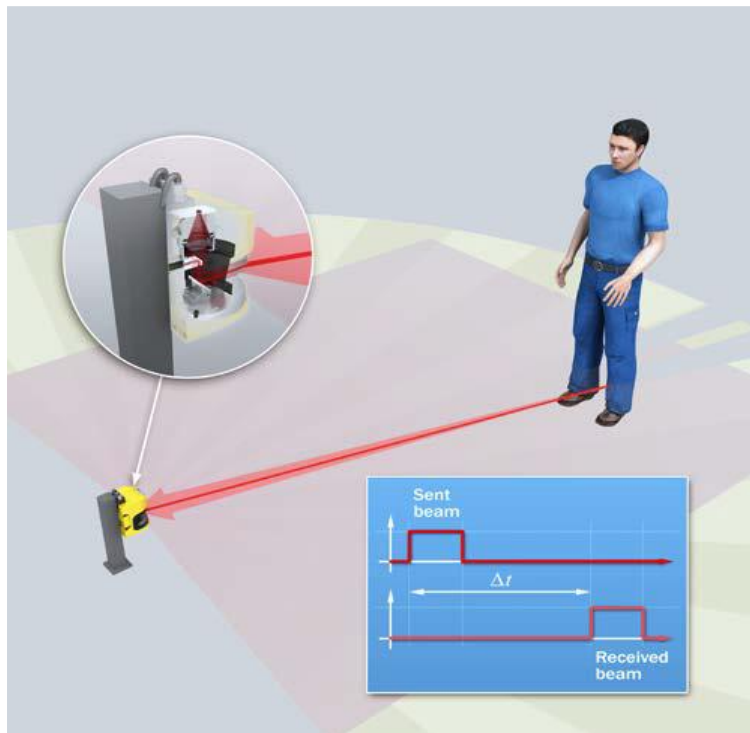


Figura 14. Modelo de detecção de *scanner laser* de segurança [24].

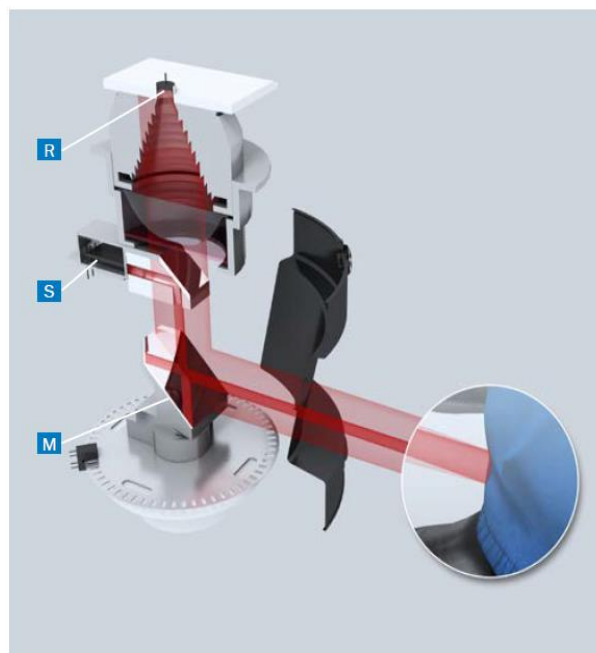


Figura 15. Estrutura básica de um *scanner a laser* [24].

Os *scanners a laser* de segurança usam pulsos de luz individuais em direções precisas e não cobrem continuamente a área a ser monitorizada. Através deste princípio de funcionamento são obtidas resoluções (capacidades de detecção) entre 30 mm e 150 mm. Com o princípio de varredura ativa, os *scanners a laser* de segurança não precisam de recetores ou refletores externos. Estes devem detetar de forma fiável objetos com refletividade extremamente baixa (por exemplo, vestuário de trabalho preto). A norma internacional IEC 61496-3 [17] estabelece os requisitos de segurança para AOPDDR.

- **Dispositivos de proteção à base de câmera (VBPD)**

Os VBPDs são dispositivos de proteção baseados em câmera e usam tecnologias de captura e processamento de imagens para a detecção de segurança de pessoas (Figura 16). Os emissores especiais de luz são atualmente utilizados como fontes de luz. VBPDs que usam a luz ambiente também são possíveis.



Figura 16. Proteção com câmera de segurança em célula de robô [24].

Vários princípios podem ser usados para detetar pessoas, incluindo:

- Interrupção da luz refletida a partir de um retrorefletor;
- Medição do tempo de voo da luz refletida por um objeto;
- Medição do tamanho e distância a um objeto;
- Monitoração de mudanças de padrões de fundo;
- Detecção de pessoas com base em características humanas.

2.6 Aplicação de imagem e vídeo 3D na indústria

Ainda são bastantes escassas as aplicações de monitorização através de imagem e vídeo 3D na indústria. Tal facto deve-se às características específicas das câmeras para este fim, bem como ao seu custo de aquisição. Apesar disso a Pilz, empresa de renome ao nível de segurança industrial, foi pioneira no lançamento de um sistema de monitorização 3D.

2.6.1 Sistema SafetyEYE (Pilz)

Como já foi referido, a Pilz apresentou o seu primeiro produto de monitorização e controlo espacial 3D (Figura 17).



Figura 17. Sistema de segurança 3D SafetyEYE da Pilz [26].

Este sistema de câmeras permite a monitorização e o controlo de áreas de perigo em três dimensões. Mais especificamente, deteta e sinaliza a entrada de objetos em áreas de alerta e proteção, as quais podem ser definidas livremente. O SafetyEYE determina se existem indivíduos na área de ação de movimento perigoso (proteção) ou se estes entraram numa zona de nível de segurança elevado (segurança).

Na Figura 18, podemos visualizar o sistema SafetyEYE, aplicado a uma máquina, estando este no seu topo, de forma a que não exista qualquer obstáculo entre as áreas definidas pela câmera e a posição do operador.

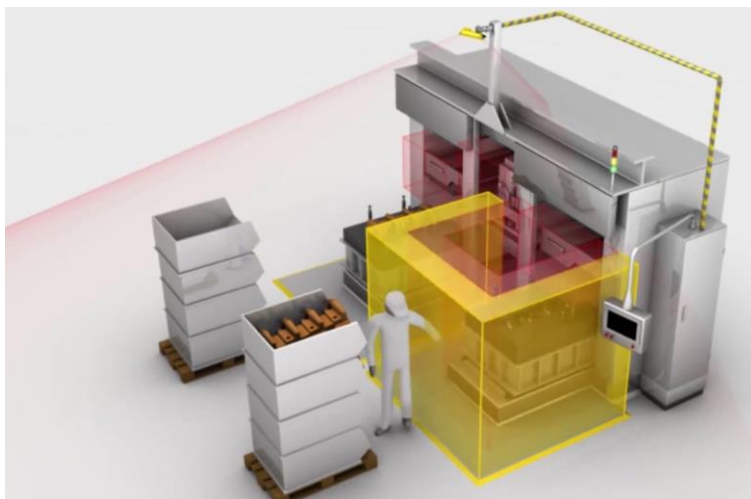


Figura 18. Áreas de proteção do operador, aplicadas a uma máquina [27].

Com este sistema de segurança, é possível monitorizar sem interrupção as áreas de trabalho perigosas, especialmente em combinação com robôs (Figura 19), definindo a sua área de trabalho, de modo a que o operador não se insira no ciclo de trabalho do robô.

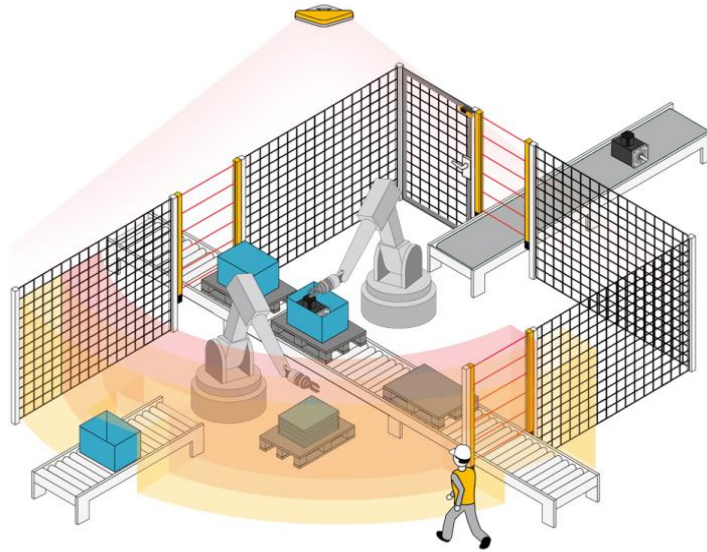


Figura 19. Área de trabalho de um robô numa linha de montagem [27].

As áreas de proteção virtuais podem ser criadas com auxílio do configurador do SafetyEYE (Figura 20). Além de definir as áreas de proteção, é possível agrupar ou alterar as disposições da área de proteção de acordo com as necessidades e redefinir a disposição com o configurador do SafetyEYE através do rato.

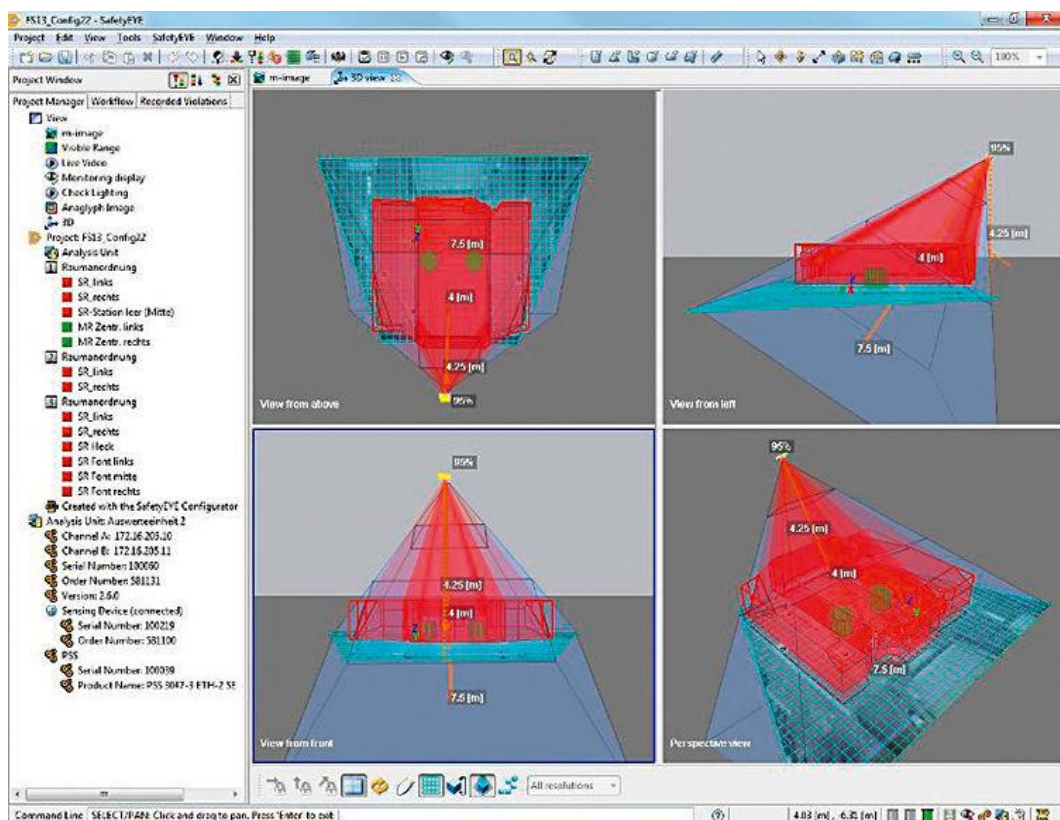


Figura 20. Criação de áreas de proteção através do configurador do SafetyEYE [28].

Este sistema apresenta características bastante interessantes, nomeadamente a sua própria iluminação, bem como a proteção IP 65. No que se refere às dimensões de abrangência da câmara, as suas características encontram-se especificadas na Figura 21. Obedece, ainda, a todas as normas dentro da categoria na qual está inserido. Mais especificamente apresenta as seguintes características:

- Resolução de braço até uma altura de instalação de 4 metros;
- Proteção do corpo;
- Máximo campo visual é de aproximadamente 72 m²;
- Iluminação requerida a partir de 300 Lux, dependendo do fundo e do ambiente;
- Projetado para aplicações até Cat. 3 de acordo com a EN 13849-1:2008 [19], SIL2 de acordo com a IEC 61508 [22], PL d de acordo com a EN ISO 13849-1 [19] e aplicações de acordo com a DIN EN 61496 [16];
- Tipos de proteção: IP65 (unidade de sensores), IP20 (unidade de avaliação).

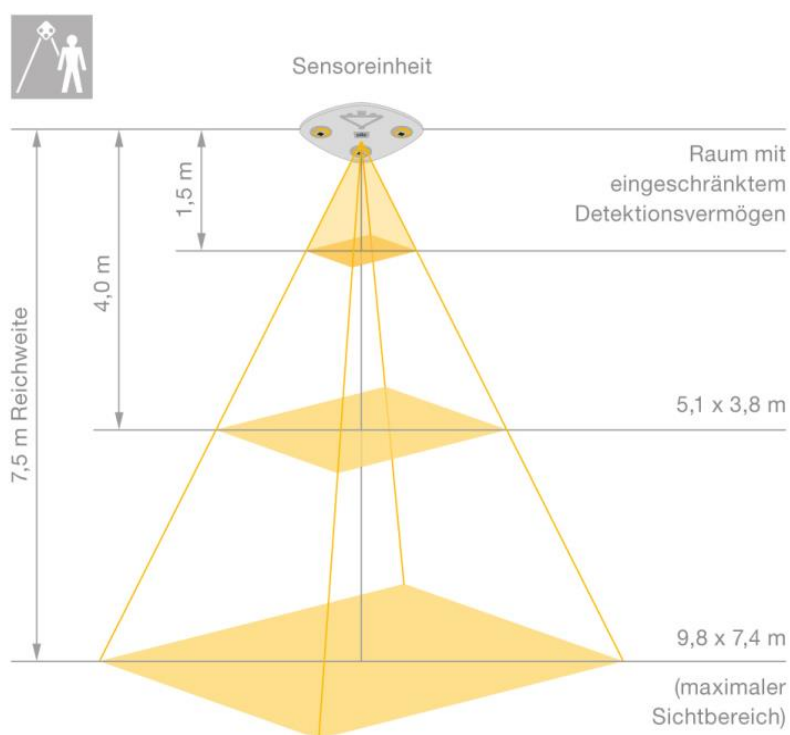


Figura 21. Demonstração das áreas abrangidas pelo SafetyEYE [26].

2.7 Outras áreas de aplicação de imagem e vídeo 3D

A utilização de imagem e vídeo 3D pode ser útil nas mais variadas áreas, apesar de ser na indústria que esta tecnologia se começa a instalar. A sua adaptação a novos ambientes é na verdade bastante flexível, desde que os parâmetros principais, tais como, a luminosidade e as condições de reflexão, se mantenham em boas condições. Desta forma, existe um conjunto elevado de aplicações de monitorização 3D possíveis de construir, indo ao encontro dos padrões de segurança.

2.7.1 Sistema de monitorização de zona 3D QuadCam

O QuadCam (Figura 22) é um sistema de controlo e deteção humana baseado na tecnologia de visão 3D para monitorizar um volume previamente definido no espaço. A zona de deteção é definida de forma detalhada e consiste numa região perimetral e numa área central interna. Quando o volume 3D definido é invadido, o sistema QuadCam envia um sinal, que pode ser usado para ativar uma variedade de dispositivos.



Figura 22. Sistema QuadCam [29].

O QuadCam deteta um indivíduo ou indivíduos que entrem nas zonas de deteção, usando *software* de imagem 3D. Distingue entre um, dois ou mais indivíduos, proporcionando monitorização em tempo real de pessoas dentro de seu campo de visão. Este sistema ignora também objetos de fundo, concentrando-se em indivíduos ou objetos indesejados detetados no seu campo de visão.

O sistema, desenvolvido e fabricado pela empresa Newton Labs, é composto por um sensor de quatro câmaras ligado a uma unidade de controlo. Cada câmara é capaz de monitorizar uma área de 2,4 m x 3 m (Figura 23), o que se reflete numa pequena área de 7,2 m². Contudo, o QuadCam permite a monitorização de 10 áreas separadas ou, de forma semelhante, as imagens podem ser sincronizadas para monitorizar uma única área. A zona é definida pelo utilizador, que desenha as zonas com a interface do utilizador do PC, independentemente da forma. O sistema possui um tempo de reação rápido, enviando um sinal dentro de 100 milissegundos após uma invasão.

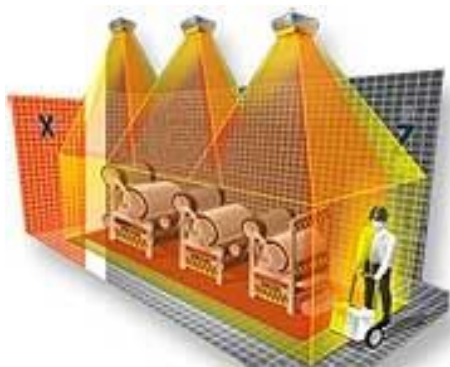


Figura 23. Monitorização de áreas com o sistema QuadCam [29].

Como sistema de deteção e monitorização baseado em zonas, o QuadCam é bastante adaptável. Pode, por exemplo, ser ligado a um sistema de alarme, a fim de estabelecer uma barreira de

segurança da área interdita numa galeria de arte ou museu (Figura 24). No ambiente industrial, o QuadCam pode por exemplo ser ligado a máquinas em células de trabalho sem barreiras, para criar uma zona de segurança definida com precisão, como uma região de alarme e uma região de paragem da máquina.



Figura 24. Possível área de monitorização com o sistema QuadCam [29].

3 Aquisição e modelização 3D com sensores RGB-D

A aquisição de imagens e vídeo com sensores RGB-D tem tido recentemente um crescimento muito acelerado. Tal facto deve-se principalmente aos sensores e tecnologias de aquisição que foram desenvolvidos nos últimos anos, e que permitiram melhorar muito significativamente a exatidão e precisão das medidas. Neste capítulo, introduzem-se algumas dessas tecnologias de aquisição com sensores RGB-D, incluindo uma apresentação dos conceitos que servem de base a essas tecnologias, focando mais especificamente a tecnologia TOF.

3.1 Geometria e captura de imagem e vídeo 3D

Existem atualmente diferentes técnicas para a obtenção de descrições completas de ambientes 3D ou apenas da informação associada à profundidade dos objetos. A maioria dos métodos de mapeamento 3D pode ser globalmente classificada em métodos ativos e métodos passivos (Figura 25). Dentro dos métodos ativos, as técnicas envolvem tipicamente uma forma de iluminação específica cuidadosamente controlada para obtenção da informação 3D, enquanto as técnicas associadas aos métodos passivos utilizam tipicamente iluminação comum, como por exemplo a luz ambiente. As técnicas ativas são geralmente mais simples em termos do processamento necessário para obter a informação 3D, pois a iluminação pode ser controlada com esse objetivo específico. As técnicas passivas podem, por exemplo, basear-se em imagens de um conjunto de vistas diferentes, a partir das quais as imagens são adquiridas e/ou iluminadas.

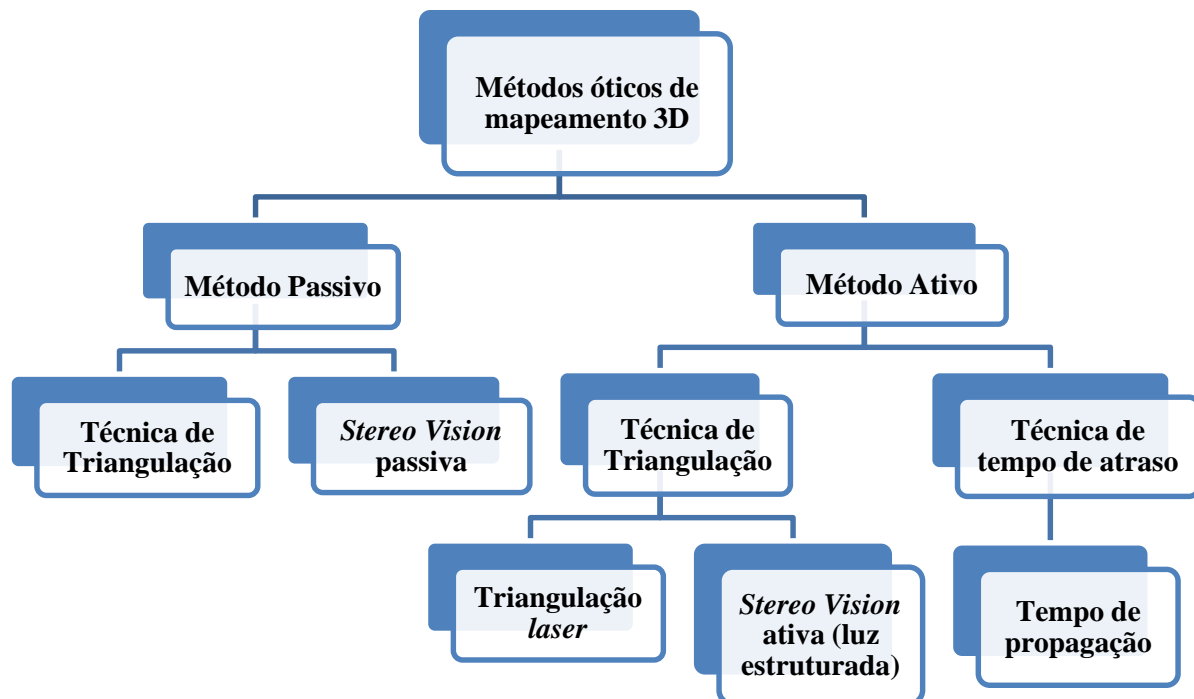


Figura 25. Relação dos métodos óticos em 3D, com base em [30].

As técnicas ativas dividem-se em técnicas de triangulação ativa e em técnicas de tempo de atraso ou tempo de propagação. A triangulação é tipicamente implementada utilizando luz *laser*

ou através da projeção nos objetos de um padrão de luz estruturada na gama dos infravermelhos ou *InfraRed* (IR). As técnicas que utilizam díodos emissores de luz ou *Light Emitting Diodes* (LED) a operar na gama dos infravermelhos, possuem a vantagem importante de ocorrerem muito menos interferências provocadas pelas condições de iluminação ambiente que são, geralmente, muito variáveis.

As técnicas passivas funcionam apenas com a luz do meio envolvente, e são habitualmente divididas em técnicas de triangulação e técnicas de visão estéreo passiva. Por esta razão, os diversos comprimentos de onda presentes na radiação luminosa e os vários tipos de reflexão que podem ocorrer, podem comprometer a fidelidade da informação 3D recolhida para pequenas áreas. Têm, ainda assim, a vantagem relativamente às técnicas ativas, de não ser necessário nenhum tipo de iluminação artificial, minimizando a complexidade dos sistemas.

3.2 Técnicas de aquisição de imagem 3D

Existem diversas técnicas desenvolvidas para a aquisição de imagem e vídeo 3D. Nas seções seguintes apresentam-se as que se consideram mais relevantes, tendo em conta os objetivos específicos do projeto proposto no Capítulo 4.

3.2.1 *Stereo Vision*

Stereo Vision é a técnica de aquisição de imagem 3D mais conhecida e utilizada, sendo alvo de constante pesquisa e de novos desenvolvimentos. É também uma técnica instalada nas mais variadas áreas da indústria. Esta técnica tem como base o funcionamento do sistema visual humano. A informação de profundidade é obtida usando duas ou mais imagens adquiridas por duas ou mais câmeras, deslocadas no espaço, ou apenas por uma única câmera que se movimenta em torno da cena a capturar. As imagens da direita e da esquerda são ligeiramente deslocadas e a informação de profundidade é obtida através dessa disparidade entre as imagens. Se pudermos encontrar os pontos correspondentes nas duas imagens, isto é, os pontos nas duas imagens que correspondem à projeção do mesmo ponto da cena ou objeto observado, então a profundidade pode ser calculada através da técnica designada triangulação, tal como representado na Figura 26.

Existem dois problemas principais na técnica da visão *stereo*, designadamente, o problema da correspondência e o problema da reconstrução. Encontrar os pontos correspondentes de um ponto no mundo real nas duas imagens é o problema da correspondência. Uma vez obtidos os pontos correspondentes, podemos calcular a disparidade, que é a diferença entre dois pontos na imagem, em termos de píxeis. Uma vez calculada a disparidade, podemos reconstruir o ponto 3D, se a geometria do conjunto de aquisição estéreo for conhecida.

A resolução teórica da profundidade da câmera *stereo* é influenciada pela geometria do sistema de aquisição e pela resolução da quantização de coordenadas. No entanto, a resolução de profundidade efetivamente obtida a partir de uma câmera estéreo (Figura 27) é limitada por vários fatores, tais como a imprecisão da calibração e erros de pontos de correspondência [31]. Os dois parâmetros da câmera, comprimento focal e linha de base, afetam a resolução de profundidade obtida em sistemas estereoscópios.

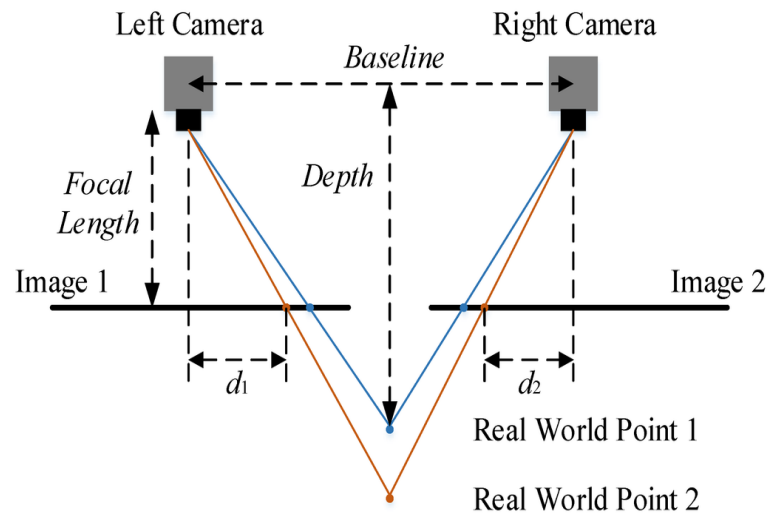


Figura 26. Método de Triangulação *Stereo Vision* [32].



Figura 27. Câmera de *Stereo Vision* [33].

3.2.2 Triangulação *laser*

Outra técnica utilizada para aquisição de informação de profundidade é a Triangulação *Laser*. O sistema de triangulação a *laser* para aquisição de informação 3D consiste em três fases principais: a transmissão, a recepção e o processamento. A transmissão consiste num díodo de *laser*, que emite a luz até ao objeto cuja profundidade deve ser medida. Essa luz refletida é recebida e direcionada para um detetor. Dependendo do detetor, o processador processa a saída de modo a fornecer a informação de distância (Figura 28). A posição da luz refletida no detetor (X') varia de acordo com a posição do objeto relativamente à referência (X). Assim, a posição da luz incidente no detetor fornece a informação de profundidade/distância ao objeto.

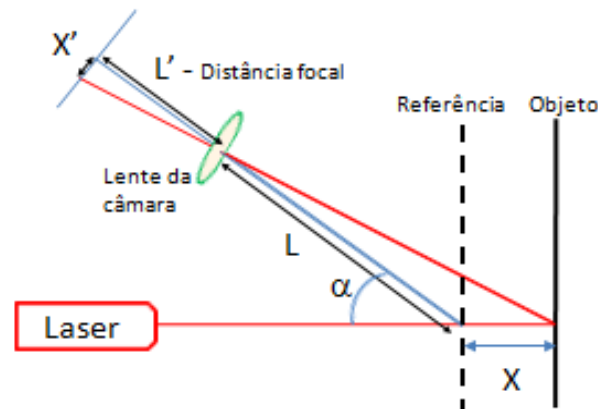


Figura 28. Método de Triangulação *laser* para aquisição de dados 3D [30].

O deslocamento X vem em função do afastamento do objeto em relação ao ponto de referência. A distância X é dada pela Equação (3.1).

$$X = \frac{X'L}{L' \sin \alpha - X' \cos \alpha} \quad (3.1)$$

Os sensores de triangulação a *laser* são de dois tipos: difusos e especulares. O sensor é selecionado com base nas propriedades de reflexão do objeto a ser medido. Um sensor especular é adequado para materiais onde se verifica predominantemente a reflexão especular, enquanto os sensores difusos são adequados para materiais onde a reflexão é difusa. Os sensores de triangulação a *laser* bem projetados oferecem uma resolução e estabilidade extremamente elevadas, muitas vezes aproximando-se dos complexos sistemas de interferómetro a *laser*. A sua capacidade de detetar movimentos muito pequenos permite que sejam utilizados com sucesso em muitas aplicações de medição, exigentes e de alta precisão. O principal fator na limitação da resolução é o ruído elétrico dos sistemas [34].

3.2.3 Time Of Flight

A técnica de tempo de propagação do sinal de luz ou *Time Of Flight* (TOF), tem sido amplamente utilizada em aplicações de Radar [35] e Lidar [36] há cerca de trinta anos [37]. Este sistema consiste, fundamentalmente, em enviar um sinal na direção de um objeto e medir as propriedades do sinal devolvido pelo objeto. Estas propriedades do sinal devolvido são usadas para medir o tempo de propagação (TOF), assim a distância a que se encontra o objeto é obtida multiplicando o tempo de propagação pela velocidade do sinal no meio físico em questão. O tempo de propagação pode ser medido modulando um sinal por impulsos ou por um sinal contínuo [38]. Os sensores baseados em impulsos medem o tempo de propagação de um impulso de luz, enquanto os sistemas baseados na modulação por um sinal contínuo medem a diferença de fase entre o sinal emitido e o sinal recebido. Na Figura 29 está representado um sistema de medição da profundidade por medição da diferença de fase proposto por Foix [38], que utiliza a modulação por um sinal contínuo que poderá ou não ser um sinal sinusoidal. Este sistema consiste na emissão de luz infravermelha através de LEDs que é refletida de volta ao

sensor, onde cada um dos pixéis do sensor retira uma amostra dessa mesma luz refletida pelo objeto quatro vezes em cada período do sinal (m_0 , m_1 , m_2 e m_3). Estas quatro quantidades permitem reconstruir o sinal recebido. O deslocamento de fase entre o sinal emitido e o sinal recebido pode ser definido pela Equação (3.2) e a distância D aos objetos pode ser obtida a partir da Equação (3.3).

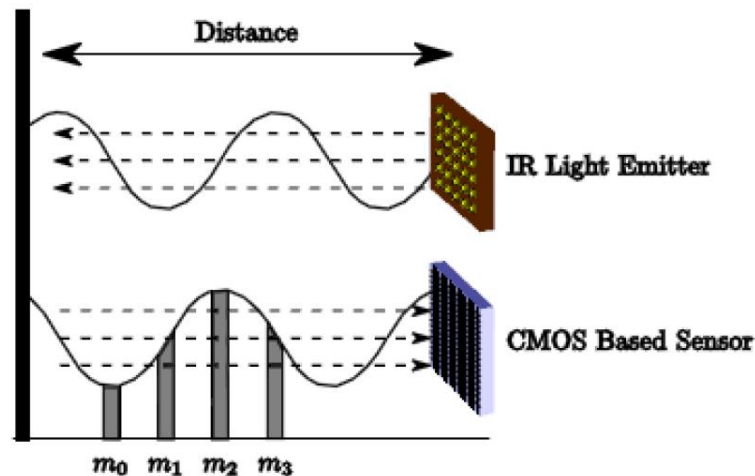


Figura 29. Medição da distância com a tecnologia TOF [38].

A diferença de fase entre o sinal emitido e o sinal recebido é:

$$\phi = \arctan\left(\frac{m_3 - m_1}{m_0 - m_2}\right) \quad (3.2)$$

A diferença ao objeto é:

$$D = L \frac{\phi}{2\pi} \quad (3.3)$$

onde L é a distância máxima permitida pelo sensor, que é uma característica relacionado com a frequência de modulação da luz emitida (f_m), e dada pela Equação (3.4),

$$L = \frac{c}{2f_m} \quad (3.4)$$

onde c representa a velocidade da luz.

Um dado importante nos sensores de detecção de profundidade é o efeito de máscara, que surge pelo facto de o sinal modulado ser periódico. Assim não será possível determinar a profundidade de objetos que estão a uma distância em que o sinal de luz devolvido tem um deslocamento de fase superior a 360 graus. É possível construir um sistema que usa frequências múltiplas como forma de eliminar este tipo de ambiguidade de acordo com Gokturk [37].

A tecnologia TOF é uma ferramenta poderosa para aplicações de robótica móvel, especialmente para operações em tempo real, uma vez que uma das mais acentuadas características dos sensores é a capacidade de processar imagens a uma velocidade de até 30 frames por segundo [39], para além de outro tipo de aplicações [40]. As câmaras baseadas na técnica de tempo de propagação são relativamente recentes e permitem definir imagens tridimensionais a uma

velocidade consideravelmente elevada, assim como permitem obter informação da profundidade para todos os pixéis.

Na Figura 30 está representado o exemplo de uma câmara comercializada pela Microsoft (Kinect V2). Nos sistemas existentes atualmente, o número de pixéis existentes nas imagens ainda é relativamente reduzido; o Kinect V2 da Microsoft tem uma matriz de 512 pixéis na horizontal e 424 pixéis na vertical. Adicionalmente, nos sistemas atuais, não é possível eliminar completamente, ao ponto de ser desprezável, o ruído presente nos valores de profundidade obtidos depois da calibração. No entanto, estes métodos de mapeamento 3D baseados em TOF estão a mostrar um enorme potencial em diversos domínios científicos [38]. Nos últimos anos, estes sistemas têm apresentado avanços significativos em termos de desempenho, nomeadamente, no aumento da resolução, na minimização dos erros e no aumento das velocidades de aquisição.



Figura 30. Câmara de tempo de propagação do sinal (TOF) [41].

3.3 Geometria e características de sensores RGB-D

Os sensores RGB-D possuem um conjunto de características técnicas de muito elevado potencial, o que torna previsível que as câmaras com base na tecnologia TOF venham progressivamente a substituir outras tecnologias, ou, alternativamente, venham a ser usadas como seu complemento, em muitas áreas de aplicação [38]. Estas características incluem a capacidade de obter informação da profundidade associada a cada pixel das imagens, elevada capacidade de cálculo nos próprios dispositivos permitindo processamento complexo, elevada velocidade de aquisição em termos do número de imagens ou frames por segundo (*fps*) e, ainda, robustez em relação à alteração das condições de luminosidade. Adicionalmente, estas características são possíveis sem a necessidade de técnicas de câmaras ou objetos móveis para aquisição de profundidade.

Na Tabela 4 apresentam-se as características mais comuns de algumas câmaras RGB-D disponíveis no mercado.

Tabela 4. Características das diferentes marcas de câmeras 3D.

	Câmera cores	Resolução	Distância máxima profundidade	Distância mínima profundidade	Visão horizontal	Visão vertical	Visão diagonal	Técnica
Kinect V1	640x480 30 fps	320x240	4,5 m	0,4 m	57°	43°	X	Luz estruturada
Kinect V2	1920x1080 30fps	512x424	8 m	0,5 m	70°	60°	X	<i>Time Of Flight</i>
Asus Xtion	SXGA (1280*1024)	VGA (640x480) 30 fps, QVGA (320x240) 60 fps	3,5 m	0,8 m	58°	45°	70°	Luz estruturada
Mora	5MP, 1.12 um, BSI (1/5")	VGA, 3um (1/7.5")	X	X	X	X	82°	Luz estruturada
Lima	X	X	X	X	X	X	90°	Luz estruturada
Zora	X	VGA, 3um (1/7.5")	X	X	X	X	70°	Luz estruturada
CamBoard Nano	90 fps	X	X	X	90°	68°	X	<i>Time Of Flight</i>

Tendo em consideração a informação na Tabela 4, facilmente se justifica a opção pela utilização da câmara Kinect V2. A negrito encontra-se a melhor qualidade do parâmetro a analisar. Através de uma rápida análise verifica-se que o Kinect V2 possui as melhores especificações para o objetivo deste projeto; por exemplo, possui um valor máximo para a profundidade que pode ir até aos 8 metros. Embora nas características originais da câmara seja indicada uma distância máxima de 4,5 metros para a profundidade, realizaram-se vários testes que permitiram concluir que, para a aplicação a desenvolver neste projeto, o sensor é utilizável até 8 metros de profundidade. Aliando-se a esta característica, junta-se ainda o facto de utilizar a tecnologia TOF, com todas as vantagens já descritas. Ainda a favor do Kinect V2, tem-se a facilidade com que se efetua a ligação a um computador, através de uma interface USB 3.0, e a rápida integração de *software*, permitida através de um SDK disponibilizado pela Microsoft.

3.4 Geometria e captura detalhada com o sensor Kinect V2

Nos últimos anos, as câmeras RGB-D têm registado uma crescente popularidade, principalmente devido ao surgimento da série Microsoft Kinect [42], que se tornou um acessório fundamental para as consolas de jogos Xbox (Microsoft). Em 2010, o Kinect V1 foi lançado utilizando a tecnologia de luz estruturada para capturar a informação de profundidade. A segunda versão do Kinect, o Kinect V2, foi lançado em 2014, juntamente com a nova versão da Xbox. Este novo Kinect V2 utiliza uma tecnologia de luz modulada e TOF para a medição de profundidade. Assim, permite obter a informação de profundidade com muito maior resolução, limitando também a interferência de fontes externas [43], [44].

Comparado com a maioria dos *scanners* 3D, o Kinect é muito acessível, o que o torna uma ferramenta bastante interessante para o ambiente de desenvolvimento, assim como para empresas interessadas em integração 3D. Contudo, o Kinect foi originalmente concebido de forma a identificar os movimentos do corpo humano [30]. Uma vez que a monitorização tem de ser feita em tempo real, o Kinect foi concebido de forma a poder capturar as frames de profundidade com frequências elevadas, sendo o máximo de 30 Hz.

O Kinect V2 fornece ainda um *Network User Interface* (NUI) que permite detetar o movimento do corpo humano, gestos, comandos falados e até reconhecimento facial. Esta *interface* permite que os utilizadores interajam diretamente com o Kinect através do computador.

3.4.1 Sensor de Profundidade

O Kinect V2 tem três fontes de luz infravermelha, gerando cada uma delas uma onda modulada com amplitude diferente. Para capturar as ondas refletidas, o Kinect V2 possui também uma câmara sensível à radiação infravermelha. As localizações dos *lasers* emissores de luz e dos sensores para a radiação infravermelha são mostradas na Figura 31.

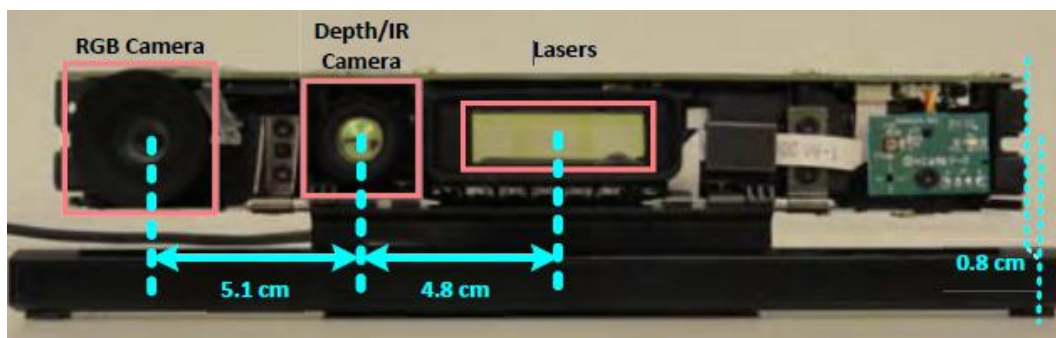


Figura 31. Localização dos *lasers* e sensores do Kinect V2 [45].

Especificações dos sensores

O sensor de infravermelhos do Kinect V2 é constituído por uma matriz de 512 por 424 *pixels* de tecnologia CMOS e de modo diferencial. As características do sensor são apresentadas na Figura 32. Cada sensor de pixel é constituído por dois fotodíodos (A, B) que são controlados pelo mesmo sinal de relógio que também controla a modulação da onda emitida. Os fotodíodos convertem a luz capturada em corrente que pode ser medida. Os díodos são acionados pelo relógio de modo que se $A = [a_i]$ é ligado, $B = [b_i]$ é desligado, e vice-versa. Estando assim, de acordo com o descrito [45]:

- $([a_i] - [b_i])$ mostra a correlação entre a luz recuperada e o sinal de relógio e pode ser usado para obter informação de fase (imagem de profundidade);
- $([a_i] + [b_i])$ fornece uma imagem em escala de cinza, regular, iluminada pela luz ambiente normal (imagem ambiente);
- $\sqrt{\sum_i ([a_i] - [b_i])^2}$ fornece imagem em escala de cinza que é independente da iluminação ambiente (imagem ativa).

Feature	Kinect for Windows 1	Kinect for Windows 2
Color Camera	640 x 480 @30 fps	1920 x 1080 @30 fps
Depth Camera	320 x 240	512 x 424
Max Depth Distance	~4.5 M	~4.5 M
Min Depth Distance	40 cm in near mode	50 cm
Horizontal Field of View	57 degrees	70 degrees
Vertical Field of View	43 degrees	60 degrees
Tilt Motor	yes	no
Skeleton Joints Defined	20 joints	26 joints
Full Skeletons Tracked	2	6
USB Standard	2.0	3.0
Supported OS	Win 7, Win 8	Win 8
Price	\$299	TBD

Figura 32. Parâmetros do sensor de profundidade [46].

O Kinect V2 dispõe ainda de um filtro passa banda que é utilizado para bloquear toda a radiação luminosa, com exceção daquela na gama de comprimentos de onda em torno de 860 nm que corresponde ao comprimento de onda do sistema de iluminação infravermelha da câmera.

O Kinect V2 também usa um mecanismo de vários obturadores que permite receber sinais a partir de vários obturadores e escolher o valor mais adequado para cada pixel. O tempo de obturação mais longo que não causar saturação do sensor é o tempo utilizado. O sistema também normaliza todos os valores relativos ao tempo de obturador mais longo.

Distância de medição

Como já referido, o Kinect V2 utiliza a tecnologia de TOF para medir distâncias [45], [47], [48]. O princípio de operação no dispositivo TOF é baseado na medição do tempo que é necessário para a onda de luz viajar do emissor ao objeto e voltar ao sensor. Seja d a distância da fonte de luz ao sensor (Figura 33), então o caso mais simples pode ser expresso como:

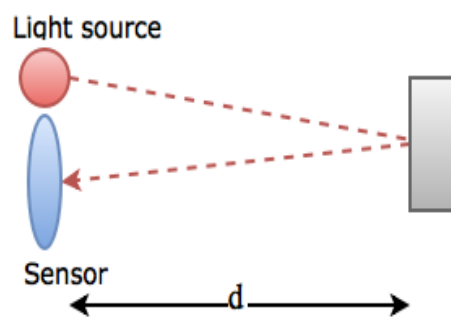


Figura 33. Princípio de funcionamento da tecnologia TOF [47].

$$d = \frac{t_r - t_e}{2} \times c, \quad (3.5)$$

onde t_e e t_r representam os instantes de tempo em que o pulso de luz é emitido e recebido respectivamente, e c é a velocidade da luz no ar. A medição de tempos associados a pulsos de luz simples é complexa e não é prática para dispositivos de captura em ambiente natural como o Kinect V2. Assim, a luz infravermelha modulada em amplitude é usada em conjunto com a recepção através do sensor de matriz CMOS. A distância é calculada com base na diferença de fase entre a onda de luz emitida e a onda de luz detetada refletida pelo objeto. A onda transmitida modulada em amplitude é:

$$T(t) = \sin(\omega t) \quad (3.6)$$

Esta onda é modulada com frequência de modulação $\omega = 2\pi f$. Então, uma distância percorrida pela onda modulada de $2d$ produz o deslocamento de fase ϕ . A Onda recebida é,

$$R(t) = \beta \sin(\omega t - \phi) \quad (3.7)$$

onde a amplitude recebida é β , que é diferente da amplitude da onda emitida e que depende de muitos fatores. No entanto, o valor absoluto da amplitude da onda recebida não é necessário

para medir distância, desde que este seja suficiente para o bom funcionamento do sistema receptor, podendo por isso ser ignorado. O deslocamento de fase depende da diferença de tempo medida:

$$\Phi = \omega t_r - \omega t_e = \omega(t_r - t_e). \quad (3.8)$$

Substituindo na Equação 3.5 resulta:

$$\Phi = \frac{2d}{c} \times \omega, \quad (3.9)$$

o que permite obter d ,

$$d = x = \frac{\Phi c}{2\omega} \quad (3.10)$$

Para obter o desvio de fase Φ , são necessárias pelo menos duas medidas do sinal recebido (uma vez que existem duas incógnitas na Equação 3.7). Para isso, o sinal recebido $\beta \sin(\omega t - \Phi)$ é misturado (por multiplicação) com uma versão do sinal de referência com um atraso de fase.

$$R_0(t) = \sin(\omega t - \Phi_{offset}). \quad (3.11)$$

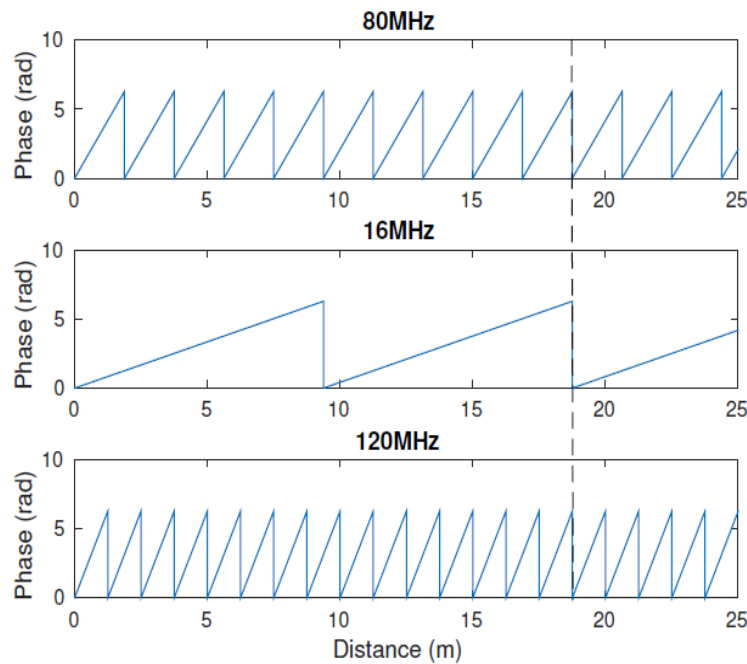


Figura 34. Relação fase-distância para as três frequências de modulação usadas [47],[49].

Esta mistura resulta em:

$$\begin{aligned} R(t) \times R_0(t) &= \beta \sin(\omega t - \Phi) \times \sin(\omega t - \Phi_{offset}) \\ &= 0,5\beta [\cos(\omega t - \Phi - \omega t + \Phi_{offset}) - \cos(\omega t - \Phi + \omega t - \Phi_{offset})] \\ &= 0,5\beta \cos(\Phi - \Phi_{offset}) - 0,5\beta \cos(2\omega t - \Phi - \Phi_{offset}) \end{aligned} \quad (3.12)$$

O resultado é processado por um filtro passa baixo, que remove a segunda parcela da adição, resultando em,

$$LP [R(t) \times R_0(t)] = 0,5\beta \cos(\phi - \phi_{offset}). \quad (3.13)$$

Usando valores diferentes para ϕ_{offset} , o desvio de fase ϕ pode ser estimado independente da amplitude β do sinal recebido. O sensor Kinect V2 utiliza três diferentes fases de *offset*, nomeadamente 0° , 120° e 240° [50].

Uma vez que a medição da distância é baseada no desvio de fase da onda modulada, a máxima distância que pode ser medida individualmente depende do comprimento de onda da onda modulada. A fase repete-se em torno de 360° (2π) criando ambiguidade para fases superiores a 2π . Isto implica que a utilização de comprimentos de onda mais longos permita medir distâncias também mais longas. Por outro lado, a utilização de comprimentos de onda mais curtos permite uma maior resolução [48]. Para permitir uma resolução elevada e simultaneamente medir distâncias mais longas, o Kinect V2 utiliza três frequências diferentes para modular o sinal de luz, nomeadamente de 120 MHz, 80 MHz e 16 MHz [45]. A distância de repetição de fase comum para estas frequências ocorre em 18,75 metros, que é também a distância máxima onde a profundidade pode ser identificada de forma única. A Figura 34 mostra a relação entre fase e distância para as frequências usadas.

Dadas três frequências $f_0 = 80$ MHz, $f_1 = 16$ MHz, $f_2 = 120$ MHz e os respetivos desvios de fase ϕ_0 , ϕ_1 , ϕ_2 , de acordo com a Equação (3.10), a distância d deve satisfazer

$$d = \frac{c}{4\pi f_0} (\phi_0 + 2\pi n_0) = \frac{c}{4\pi f_1} (\phi_1 + 2\pi n_1) = \frac{c}{4\pi f_2} (\phi_2 + 2\pi n_2), \quad (3.14)$$

onde $0 \leq n_0 \leq 9$, $0 \leq n_1 \leq 1$ e $0 \leq n_2 \leq 14$ são coeficientes de repetição de fase desconhecidos.

Substituindo para os valores das frequências usadas, vem,

$$\begin{aligned} \frac{\phi_0 + 2\pi n_0}{4\pi \cdot 10} &= \frac{\phi_1 + 2\pi n_1}{4\pi \cdot 2} = \frac{\phi_2 + 2\pi n_2}{4\pi \cdot 15} \\ \Leftrightarrow \frac{3(\phi_0 + 2\pi n_0)}{2\pi} &= \frac{15\phi_1 + 2\pi n_1}{2\pi} = \frac{2(\phi_2 + 2\pi n_2)}{2\pi} \\ \Leftrightarrow \frac{3\phi_0}{2\pi} + 3n_0 &= \frac{15\phi_1}{2\pi} + 15n_1 = \frac{2\phi_2}{2\pi} + 2n_2. \end{aligned} \quad (3.15)$$

Logo,

$$t_0 = \frac{3\phi_0}{2\pi}, t_1 = \frac{15\phi_1}{2\pi}, t_2 = \frac{2\phi_2}{2\pi} \quad (3.16)$$

O seguinte sistema de equações pode ser usado para encontrar n_0 , n_1 e n_2

$$\begin{aligned} 3n_0 - 15n_1 &= t_1 - t_0, \\ 3n_0 - 2n_2 &= t_2 - t_0, \\ 15n_1 - 2n_2 &= t_2 - t_1. \end{aligned} \quad (3.17)$$

O sensor Kinect V2 calcula e disponibiliza os coeficientes de mudança de fase, dando aos programadores a oportunidade de desenvolver os seus próprios métodos para a estimativa do mapa de profundidade real. A implementação padrão do SDK da Microsoft [51] é usada para fornecer a profundidade em milímetros para cada *pixel*.

3.4.2 Transformação de profundidade em coordenadas 3D

O sensor de TOF descrito na secção anterior fornece para cada *pixel* do sensor uma medida de distância a partir do centro geométrico da câmara. Para modelizar o espaço real e os objetos, são necessárias coordenadas 3D reais. Estas podem ser obtidas usando o modelo *pinhole* da câmara [52]. Primeiro é obtido o mapa de profundidades a partir do mapa de distâncias, e apenas depois é utilizado o modelo de câmara *pinhole* para obter as coordenadas 3D a partir do mapa de profundidades.

Distância e Profundidade

Dada a geometria na Figura 35 e o ponto genérico P da cena a capturar, podem ser identificadas as seguintes distâncias: a distância d de P ao centro da câmara C , distância focal f (que é uma característica da câmara), a distância x , do ponto principal à projeção do ponto P no plano da imagem e a distância z do centro da câmara C ao ponto P_c . A distância l entre o ponto C e a projeção de P no plano da imagem pode ser calculada usando:

$$l = \sqrt{f^2 + x^2} \quad (3.18)$$

Com base na semelhança de triângulos podemos obter:

$$\frac{z}{d} = \frac{f}{l} \quad (3.19)$$

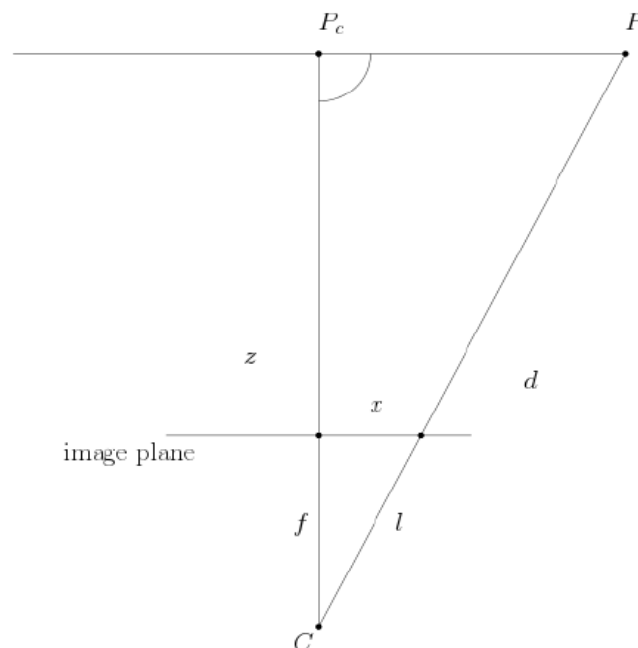


Figura 35. Cálculo da profundidade através da distância [47].

Podemos então escrever a profundidade z como,

$$z = d \frac{z}{d} = d \frac{f}{\sqrt{f^2 + x^2}} \quad (3.20)$$

O Kinect SDK disponibiliza a informação de distância já no formato de mapa de profundidades. Isso significa que, em vez da distância do centro da câmera ao ponto a considerar na cena a adquirir, o valor da profundidade associado a cada *pixel* representa a distância do centro da câmera ao plano que contém o ponto e que é perpendicular ao eixo principal da câmera.

Modelo de câmera *pinhole*

Considera-se o centro do sistema de coordenadas no centro da câmera. A Distância Focal da câmera f é a distância entre o centro da projeção e o plano da imagem. A linha que parte do centro da projeção e é perpendicular ao plano da imagem é chamada de Eixo Principal. O Ponto Principal é a intersecção entre o eixo principal e o plano da imagem.

Seja $P(X;Y;Z)$, um ponto exemplo com coordenadas 3D $(X;Y;Z)$. As coordenadas correspondentes no plano da imagem da câmera são $p(x;y)$ (nas mesmas unidades que as coordenadas 3D $(X;Y;Z)$) (Figura 36). Pela semelhança de triângulos, as coordenadas no plano da imagem podem ser escritas como:

$$x = f \frac{X}{Z}, \quad y = f \frac{Y}{Z} \quad (3.21)$$

X e Y são coordenadas reais, enquanto que na imagem as coordenadas nativas são necessariamente em *pixels*. A transformação de coordenadas requer conhecer a densidade de *pixels* por linha e por coluna (*pixels* por milímetro) no sensor de imagem. Assim k_u representa densidade de *pixels* por coluna e k_v representa densidade de *pixels* por linha. A distância entre o ponto principal e a origem de coordenadas de *pixel* também é necessária, sendo as coordenadas deste ponto $(-x_0; -y_0)$. Com esta geometria as coordenadas p do ponto no sistema de coordenadas de *pixel* podem ser escritas da seguinte forma:

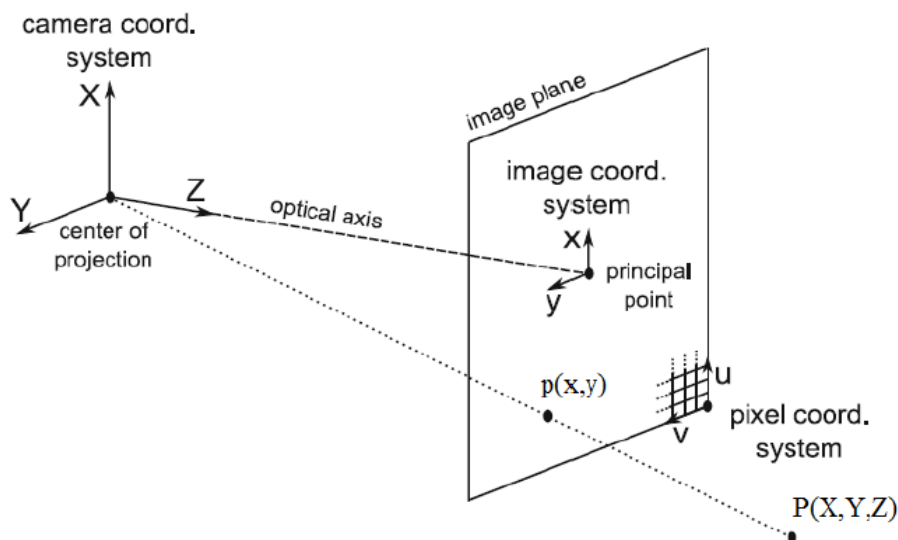


Figura 36. Modelo de câmera *pinhole* [52].

$$\begin{aligned} u &= k_u(x + x_0) = k_u f \frac{X}{Z} + k_u x_0, \\ v &= k_v(y + y_0) = k_v f \frac{Y}{Z} + k_v y_0. \end{aligned} \quad (3.22)$$

Depois de normalizar as coordenadas do ponto P , dividem-se estas pela sua coordenada Z :

$$P' = \frac{P}{Z} = \begin{pmatrix} X/Z \\ Y/Z \\ 1 \end{pmatrix} = \begin{pmatrix} X' \\ Y' \\ 1 \end{pmatrix},$$

A Equação 3.22 pode ser escrita na forma matricial,

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} k_u f & 0 & k_u x_0 \\ 0 & k_v f & k_v y_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X' \\ Y' \\ 1 \end{pmatrix} = KP' \quad (3.23)$$

Os quatro valores na matriz 3×3 são chamados parâmetros intrínsecos da câmara e são usualmente denotados como:

$$\alpha_u = k_u f, \alpha_v = k_v f, u_0 = k_u x_0, v_0 = k_v y_0. \quad (3.24)$$

Em unidades de *pixels*, estes representam a distância focal (α_u, α_v) e as coordenadas do ponto principal (u_0, v_0). A matriz intrínseca K da câmara pode ser representada como

$$K = \begin{pmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \quad (3.25)$$

Em geral, os parâmetros intrínsecos para determinada câmara são conhecidos ou podem ser obtidos por calibração. No Kinect V2, estes parâmetros podem ser obtidos a partir do SDK [53].

Considerando que u , v , K e Z são conhecidos para o ponto P , com base na discussão anterior, as coordenadas x e y podem ser obtidas através de,

$$K^{-1} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = P' = \frac{P}{Z}, \quad (3.26)$$

De onde,

$$ZK^{-1} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \quad (3.27)$$

Como Z é conhecido, X e Y podem ser calculados para todos os valores de u e v . O SDK do Kinect V2 fornece uma tabela de valores de (γ_u, γ_v) para cada pixel (u, v) , tal que

$$\begin{aligned} X &= \gamma_u \times Z \\ Y &= \gamma_v \times Z \end{aligned} \quad (3.28)$$

Distorção Radial

Na seção anterior usou-se um modelo de câmera de *pinhole* perfeito. Numa câmera real, a lente causa alguma distorção das coordenadas dos pontos, distorção esta que pode ser modelizada no plano da imagem tal como ilustrado na Figura 37.

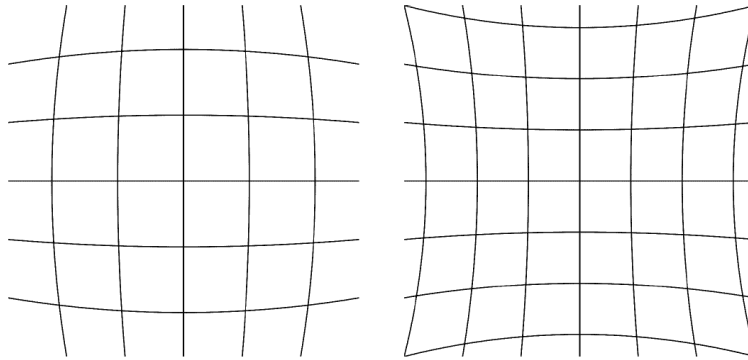


Figura 37. Distorções comuns da lente: barril (esquerda), almofada (direita) [47].

Existem muitos tipos de distorções possíveis, mas verificou-se que a distorção radial é mais comum e significativa. A maior parte da distorção pode, portanto, ser descrita pelo modelo relativo simples [54]. Sejam (x,y) as coordenadas ideais (sem distorção) no plano da imagem, e (x_d,y_d) as coordenadas observadas reais correspondentes. O ponto principal, $(0,0)$ é assumido como sendo livre de distorção sob este modelo. Então,

$$\begin{aligned} x_d &= x(1 + k_1r^2 + k_2r^4 + k_3r^6), \\ y_d &= y(1 + k_1r^2 + k_2r^4 + k_3r^6), \end{aligned} \quad (3.29)$$

onde,

$$r = \sqrt{f^2 + x^2} \quad (3.30)$$

E onde k_1, k_2, k_3 são coeficientes de distorção radial. O número de adições na Equação 3.29 não é limitado no modelo geral, porém em [54] afirma-se que dois são suficientes na maioria dos casos. O Kinect SDK fornece acesso aos três primeiros coeficientes [53].

3.4.3 Alinhamento do sensor de cor e profundidade

Uma vez que as informações de profundidade e cor são capturadas por diferentes sensores, é necessário o mapeamento entre RGB e o sensor de profundidade. A câmera RGB pode ser modelizada como uma câmera *pinhole*, da mesma forma que a câmera de profundidade descrita anteriormente. Seja K_c a matriz intrínseca da câmera para a câmera RGB, então pela Equação (3.26), pode ser escrito,

$$Z_c K_c^{-1} \begin{pmatrix} u_c \\ v_c \\ 1 \end{pmatrix} = \begin{pmatrix} X_c \\ Y_c \\ Z_c \end{pmatrix}, \quad (3.31)$$

Onde X_c , Y_c e Z_c são coordenadas de ponto P no sistema de coordenadas da câmera RGB. Como as câmeras são fixas, existe uma transformação rígida T de coordenadas de profundidade para coordenadas RGB tais que,

$$T \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} X_c \\ Y_c \\ Z_c \end{pmatrix}, \quad (3.32)$$

Usando os resultados da Equação (3.22), e da Equação (3.26), a Equação (3.33) pode ser escrita como,

$$TZK^{-1} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = Z_c K_c^{-1} \begin{pmatrix} u_c \\ v_c \\ 1 \end{pmatrix} \quad (3.33)$$

Aqui, todos os parâmetros são conhecidos ou podem ser obtidos por calibração [55]. A Equação (3.33) descreve o mapeamento entre as câmeras de profundidade e de cor. O Kinect SDK fornece funções de mapeamento convenientes em ambas as direções. É importante notar que o mapeamento se baseia no facto de que ambos os sensores poderem “ver” o ponto a ser representado. Devido a diferentes localizações dos dois sensores e também diferentes campos de visão, existem regiões onde o mapeamento direto não é possível, como pode ser observado na Figura 38.

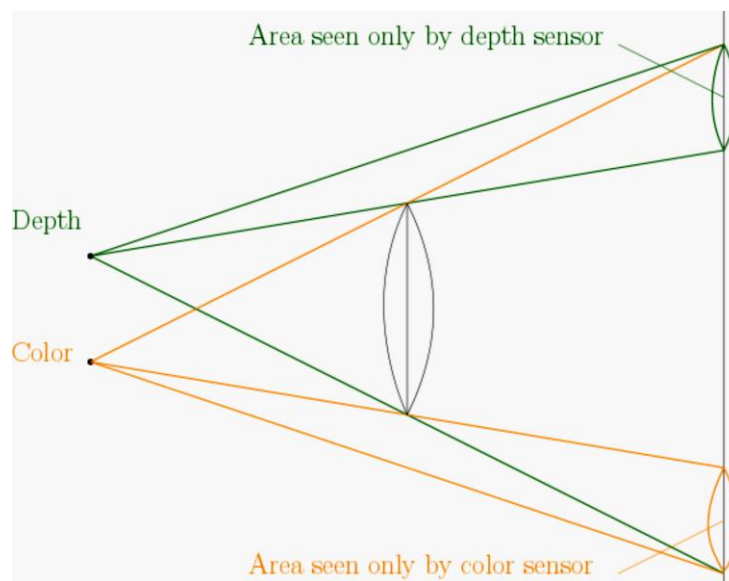


Figura 38. Profundidade e cor da câmera "pontos cegos" [47].

3.4.4 Aspectos a ter em consideração no funcionamento

A teoria descrita nas secções anteriores é perfeitamente adequada em condições ideais. Em condições reais existem vários fatores que podem causar erros na medição de profundidade. Em [50] foram analisadas várias fontes de erro e compararam-se os seus efeitos em duas gerações de Kinects.

Estudos anteriores mostraram que alguns sensores RGB-D necessitam de um tempo de pré-aquecimento antes de fornecer medições de distância confiáveis [56]. Para verificar esse fenômeno, foi determinado o tempo de espera necessário para que o Kinect V2 atinja medições mais estáveis (Figura 39). O sensor foi colocado paralelamente a uma parede plana branca e as medições de profundidade foram realizadas a cada 15 segundos durante uma hora e meia (360 mapas de profundidade registrados). Para cada mapa de profundidade, a distância do sensor à parede foi determinada em uma área central de 10 por 10 *pixels*. Esta área foi escolhida de acordo com as intensidades máximas da luz infravermelha refletida.

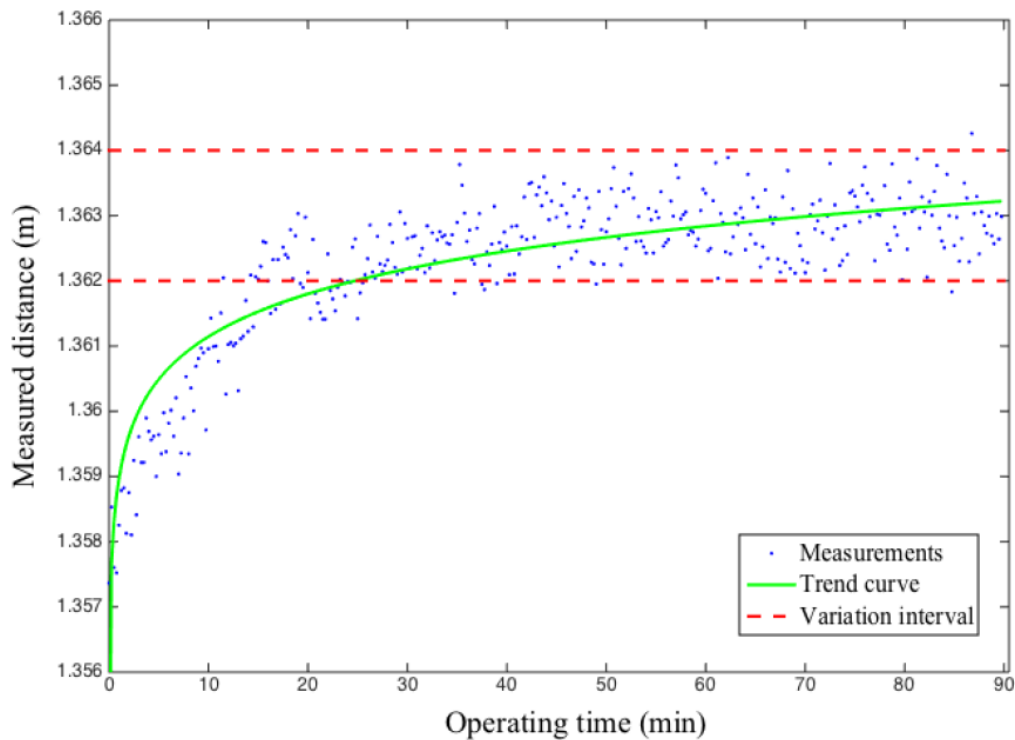


Figura 39. Valor das medições de distância em função do tempo [57].

Existe ainda o erro sistemático de distância que tem como origem as aproximações efetuadas na forma sinusoidal do sinal, as quais conduzem a erros sistemáticos da profundidade medida pelo Kinect V2. A magnitude do erro sistemático considera-se relativamente pequena, na ordem de 1-2 *mm*. Na Figura 40 podemos observar o desvio entre as distâncias reais e medidas para vários valores da distância real [57].

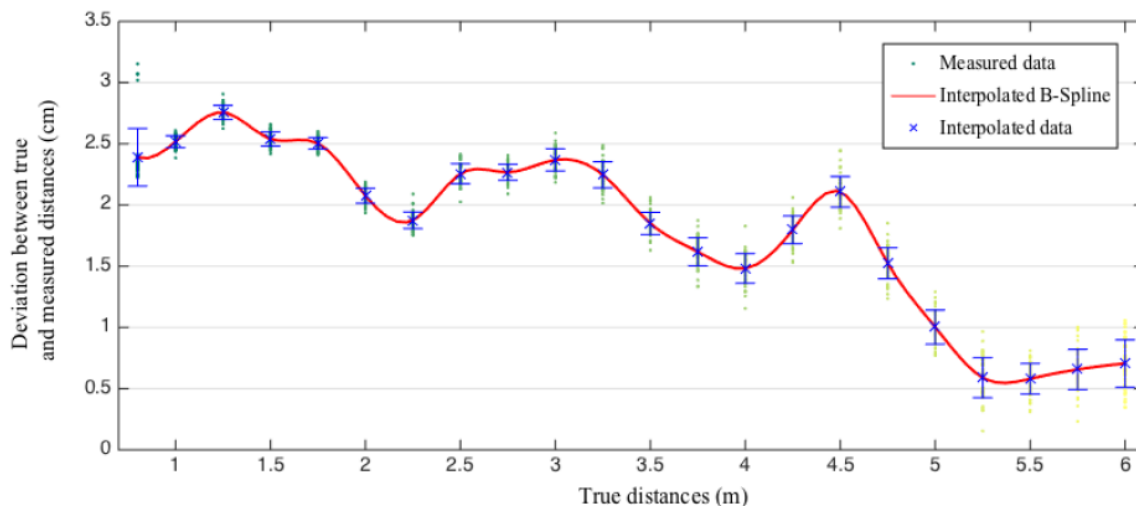


Figura 40. Desvio das distâncias medidas relativamente às medidas reais [57].

Por fim o Kinect V2 não foi concebido para aquisições ao ar livre. A influência das condições da luz natural neste ambiente pode influenciar o aparecimento de *pixels* não desejados e também diminuir o número de pontos da nuvem requerida [57].

4 Monitorização de áreas e volumes usando o Kinect V2

Neste capítulo apresenta-se o trabalho prático que foi desenvolvido neste projeto, com referência às várias fases da sua evolução. É necessário, para isso, descrever a fase de preparação do *hardware* e do *software*, passando pelo desenvolvimento de código e da sua programação e, por fim, a apresentação de resultados da fase de testes, com a qual se consolida todo o trabalho desenvolvido. Anteriormente à fase de testes em ambiente industrial, foi realizada uma fase de testes de preparação, onde foi testado o funcionamento nas várias fases da aplicação, de forma a validar o possível desempenho em ambiente industrial. Por fim, foi implementado um sistema de alarme visual e auditivo, para o utilizador perceber mais facilmente quando existe uma intrusão ou evasão da zona restrita, ou de perigo.

4.1 Especificações do projeto

O projeto tem como base o sensor Kinect V2, sendo este o *hardware* mais importante necessário para a sua implementação. Conectando-se através de USB 3.0 a um computador comum, está-se perante as condições básicas de trabalho em termos de *hardware* (Figura 41). Para se iniciar o desenvolvimento, é necessário instalar todo o *software* necessário, mais especificamente o SDK do Kinect [51], o Visual Studio [58] e o Matlab [59] com uma versão superior à 2015b. Tendo o *hardware* e o *software* anteriormente descrito conectado/instalado, estão reunidas as condições de desenvolvimento da aplicação.

De modo a ser mais perceptível para o utilizador quando existe uma invasão ou evasão da zona de perigo ou da zona interdita, procedeu-se à montagem de um circuito que integra um alarme sonoro e um alarme visual, que emite um alerta sempre que uma das zonas do espaço a ser monitorizado é infringida. Existe, por fim, um relatório de ocorrências, que especifica quando é que a zona interdita foi infringida, anotando a hora e o dia do acontecimento.

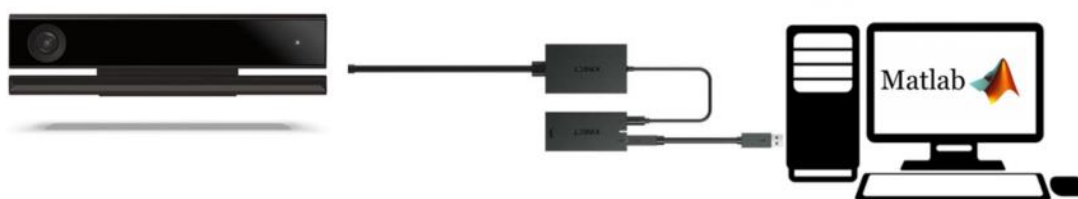


Figura 41. Ilustração da conexão do Kinect V2 ao computador.

4.2 Fase de calibração do volume interdito

A fase de calibração é fundamental para o sucesso na detecção correta de intrusão ou evasão, no/do volume monitorizado. Uma calibração incorrecta do volume pretendido por parte do utilizador, tem por consequência que a monitorização não seja efetuada na zona pretendida. De forma a simplificar o processo, e a tornar intuitiva a definição do volume interdito, a calibração é efetuada através de pontos introduzidos pelo utilizador, no espaço real, que posteriormente são mapeados no espaço virtual (*PointCloud*) de forma a serem visualizados pelo operador e monitorizados pelo sistema de detecção. Numa primeira interação com a aplicação desenvolvida, para que o utilizador tenha a perceção das distâncias da câmara à área que quer monitorizar, e posicionar melhor a câmara se necessário, é apresentada uma imagem que informa, através de uma barra lateral colorida, a distância a que os vários objetos se encontram da câmara, como se pode ver pela Figura 42.

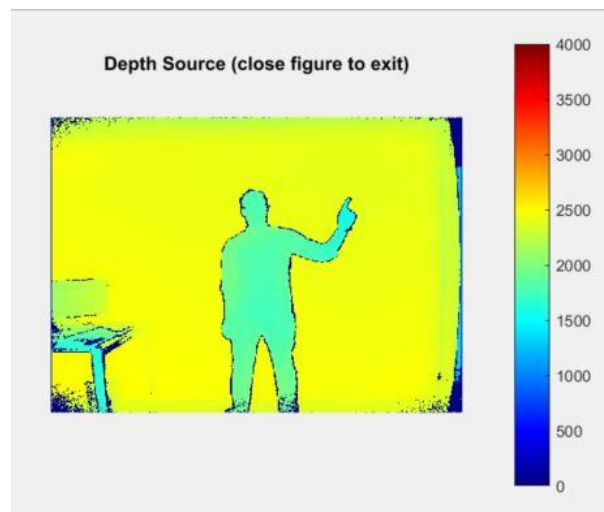


Figura 42. Distância dos objetos à câmara.

Posteriormente à visualização das distâncias numa imagem em 2D, obtém-se uma *PointCloud* do volume de visualização da câmara, através da qual é possível uma interação entre o espaço real e o espaço virtual, permitindo assim fazer a calibração do espaço 3D que se pretende que seja interdito (Figura 43).



Figura 43. Método de calibração definido pelo utilizador.

O método de calibração consiste na marcação no espaço real de 8 pontos por parte do utilizador, os quais definem os vértices para a formação do volume interdito. Marcados os pontos de A a H (adquiridos individualmente, a partir da posição do dedo indicador), o *software* guarda as suas coordenadas 3D e utiliza-as posteriormente para criar a delimitação do volume interdito. A Figura 44 mostra a criação de um volume interdito.

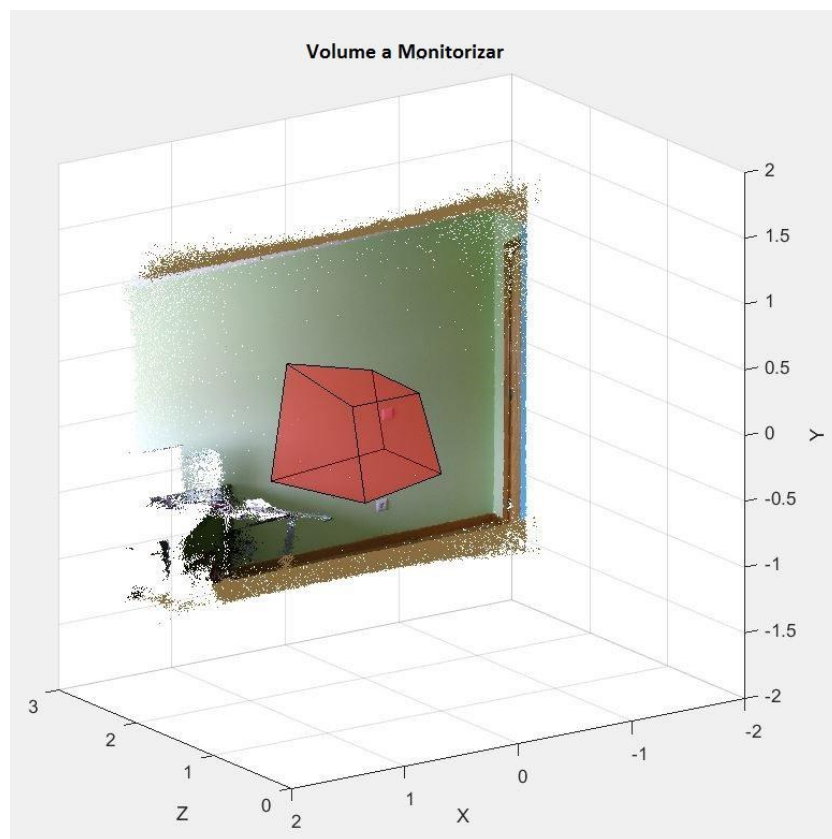


Figura 44. Volume interdito definido pelo utilizador.

O volume interdito é o espaço onde não deve existir uma invasão ou evasão. Para evitar que uma aproximação rápida resulte numa invasão deste espaço sem dar tempo para acionar os mecanismos de segurança (por exemplo parar uma máquina), criou-se uma zona de perigo (de pré-aviso ou *warning*), em torno da zona interdita, como se pode verificar pela Figura 45. Esta zona de perigo é definida automaticamente a partir dos pontos de calibração previamente definidos e da indicação da sua dimensão, que pode ser ajustável de acordo com as necessidades de uma aplicação concreta.

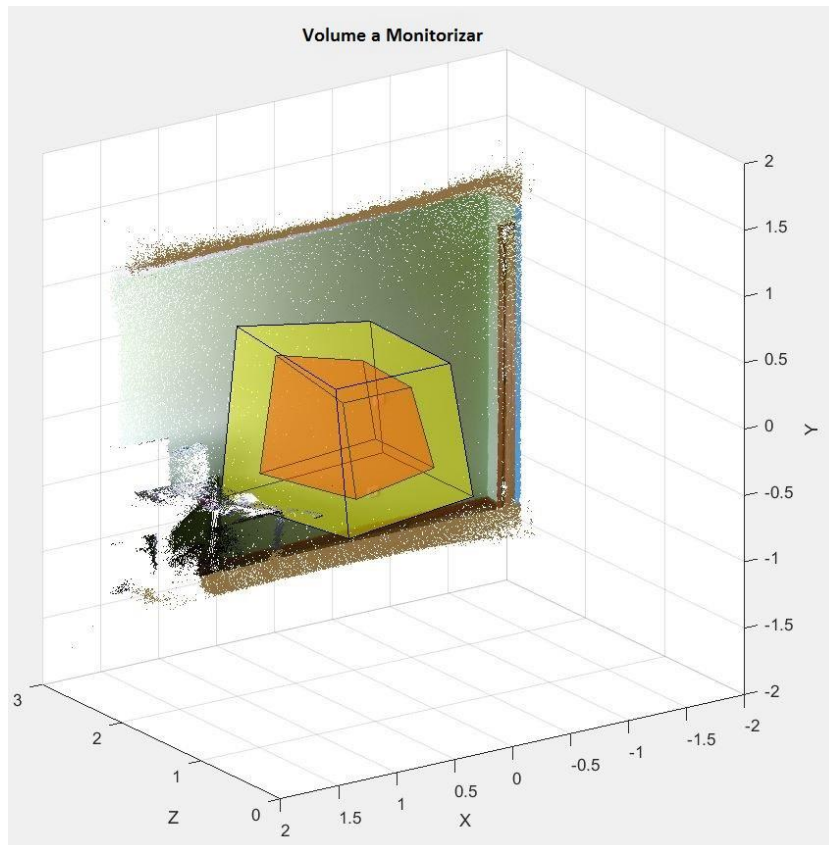


Figura 45. Zona de perigo em torno do espaço interdito.

Como forma de descrever o funcionamento da fase de calibração em termos de *software*, apresenta-se na Figura 46 um fluxograma ilustrativo, tendo em consideração as várias sub-fases do processo de calibração.

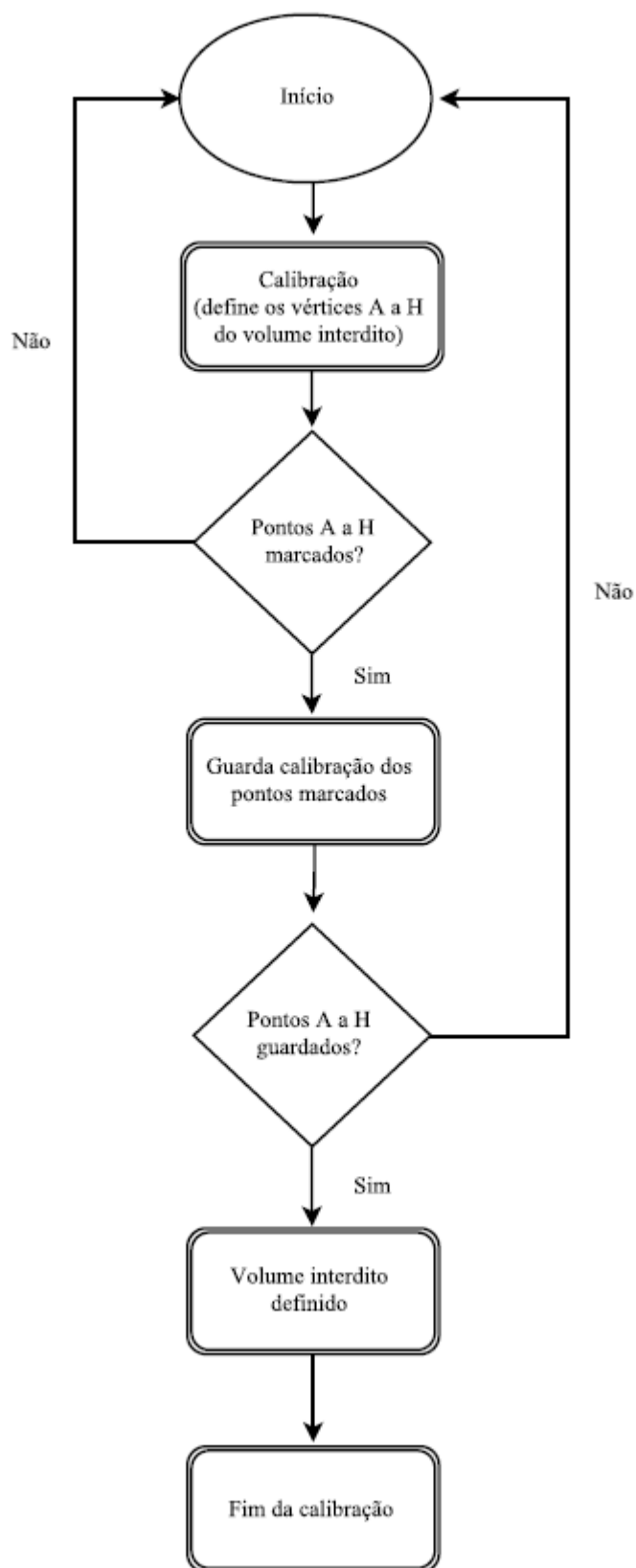


Figura 46. Fluxograma da fase de calibração.

4.3 Fase de monitorização

Adquiridos os principais pontos de calibração (vértices do volume interdito), o sistema define automaticamente o volume interdito, tendo como objetivo a sua monitorização. Para efetuar a monitorização do volume é necessário recorrer a uma técnica de divisão do espaço volumétrico de interesse em tetraedros. Após essa divisão e a quantificação do volume dos objectos que fazem naturalmente parte do volume interdito, o sistema está pronto a operar, detetando intrusões ou evasões, sempre que as mesmas se verificarem.

4.3.1 Método de Triangulação de Delaunay

As Triangulações de Delaunay são amplamente utilizadas na computação científica em diversas aplicações. Embora existam inúmeros algoritmos para calcular as triangulações, são as propriedades geométricas favoráveis da triangulação de Delaunay que a tornam tão útil em muitos problemas práticos [60], [61], [62] e [63].

A propriedade fundamental é o Critério de Delaunay. No caso das triangulações 2D, é frequentemente chamado de Critério de Circunferência Vazia. Para um conjunto de pontos em 2D, uma triangulação de Delaunay desses pontos assegura que a circunferência associada a cada triângulo (circunscrita) não contém nenhum outro ponto no seu interior. Na Figura 47, a circunferência associada com T1 está vazia, ou seja, não contém qualquer ponto no seu interior. A circunferência associada com T2 também está vazia, e como é visível não contém qualquer ponto no seu interior. Esta triangulação é uma Triangulação de Delaunay.

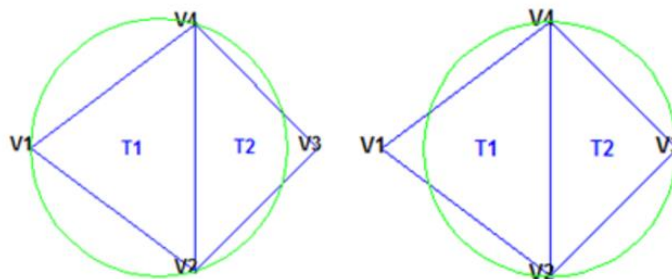


Figura 47. Circunferências com Critério de Delaunay [60].

Os triângulos da Figura 48 são diferentes, na medida em que a circunferência associada a T1 contém V3 no seu interior e a circunferência associada a T2 contém V1 no seu interior. Esta triangulação não é uma Triangulação de Delaunay.

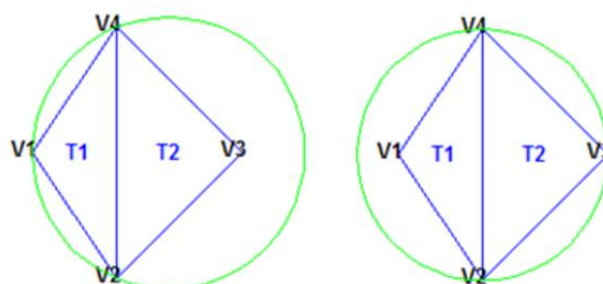


Figura 48. Circunferências com Critério de não-Delaunay [60].

Os triângulos de Delaunay são convenientemente definidos pois, ao garantir a propriedade de circunferência vazia, os triângulos com ângulos internos grandes são selecionados preferencialmente em relação aos que têm ângulos internos pequenos. Os triângulos na triangulação não-Delaunay do exemplo anterior possuem ângulos pequenos nos vértices V2 e V4. Se o lado (V2,V4) fosse substituído por um lado juntando V1 e V3, o ângulo mínimo seria maximizado e a triangulação seria desta forma uma triangulação de Delaunay.

Além disso, a triangulação de Delaunay interliga pontos numa perspectiva de vizinhos mais próximos. Ou seja, estas duas características, a criação de triângulos bem formados e a relação de vizinhos mais próximos, têm importantes implicações na prática e motivam a utilização de triangulações de Delaunay em interpolação de dados dispersos [60], [62].

As propriedades das triangulações de Delaunay estendem-se a outras dimensões. A triangulação de um conjunto de pontos 3D é composta por tetraedros. A Figura 49 mostra uma triangulação de Delaunay 3D simples, composta por dois tetraedros. A circunferência associada a um tetraedro é mostrada para realçar o critério de circunferência (esfera) vazia.

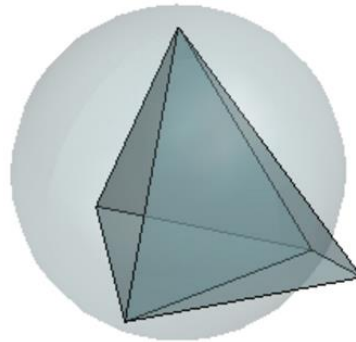


Figura 49. Triangulação de Delaunay 3D, composta por dois tetraedros [60].

4.3.2 Técnica de monitorização

A fase de monitorização engloba a obtenção, análise e tratamento dos dados fornecidos pelo sistema Kinect V2. Esta é a parte mais importante do programa desenvolvido, uma vez que é nela que os dados são obtidos e as funções de triangulação são criadas. Para a deteção atuar de forma eficaz é aplicada a técnica de Triangulação de Delaunay dentro do ciclo de monitorização. Tal como explicado na Secção 4.3.1, este método vai identificar os pontos presentes dentro do volume interdito definido pelo utilizador. Na Figura 50 pode-se visualizar o volume interdito dividido em tetraedros, após aplicada a técnica de Triangulação de Delaunay. O volume interdito é caracterizado em termos da ocupação deste volume na situação de não intrusão/não evasão. Esta caracterização é conseguida através da determinação do volume ocupado utilizando técnicas de *Point in Volume* aplicadas ao volume interdito previamente dividido em tetraedros através da Triangulação de Delaunay [63], [64] e [65].

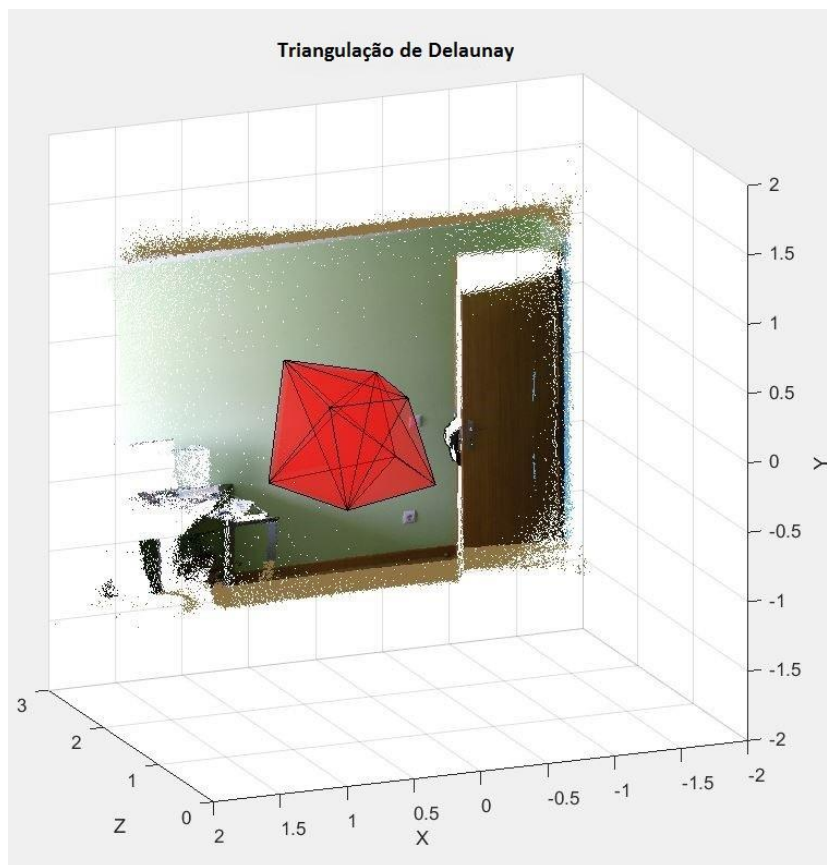


Figura 50. Divisão do volume interdito em tetraedros.

Na fase de monitorização a deteção de intrusão ou evasão é o objetivo principal da aplicação. Esta fase utiliza todas as funções descritas anteriormente e tem como objetivo informar o utilizador se houve ou não intrusão ou evasão no/do volume previamente definido. O seu funcionamento é ilustrado pelo fluxograma representado na Figura 51. O algoritmo de deteção baseia-se na comparação da ocupação de referência do volume interdito, definida na fase de calibração, com a sua ocupação calculada continuamente na fase de monitorização.

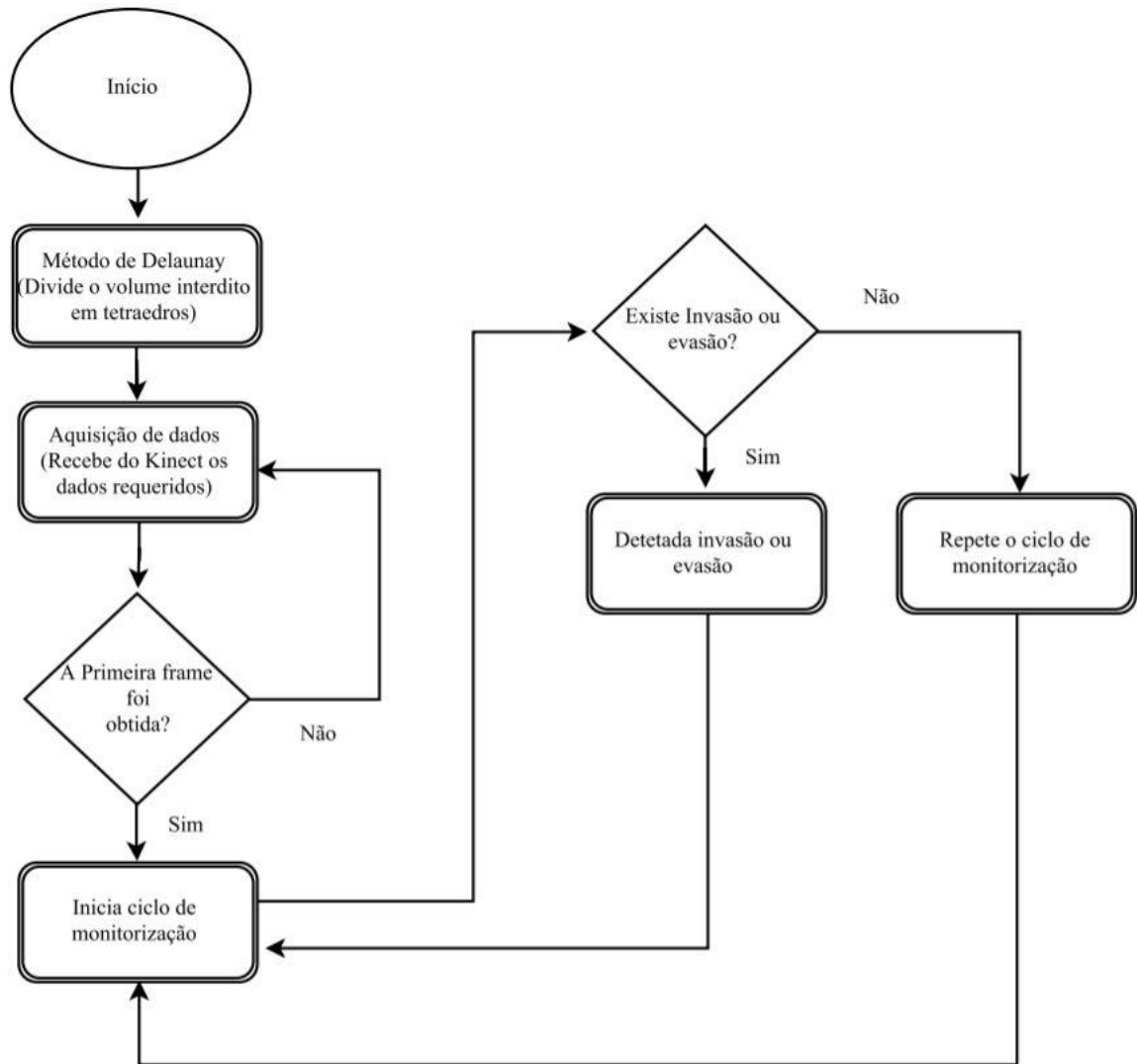


Figura 51. Fluxograma da fase de monitorização.

4.3.3 Aplicação baseada na fase de monitorização

Para validar todas as técnicas apresentadas neste Subcapítulo 4.3 procedeu-se a um teste de monitorização de um volume pré-definido na fase de calibração. Desta forma, pode-se confirmar o bom resultado da aplicação das técnicas implementadas e pode visualizar-se de uma forma intuitiva o processo de invasão da zona de perigo e da zona interdita. Apresenta-se ainda um exemplo de evasão da zona interdita.

Na Figura 52 pode observar-se o funcionamento do ciclo de monitorização, sem haver qualquer intrusão ou evasão da zona de perigo ou da zona interdita.

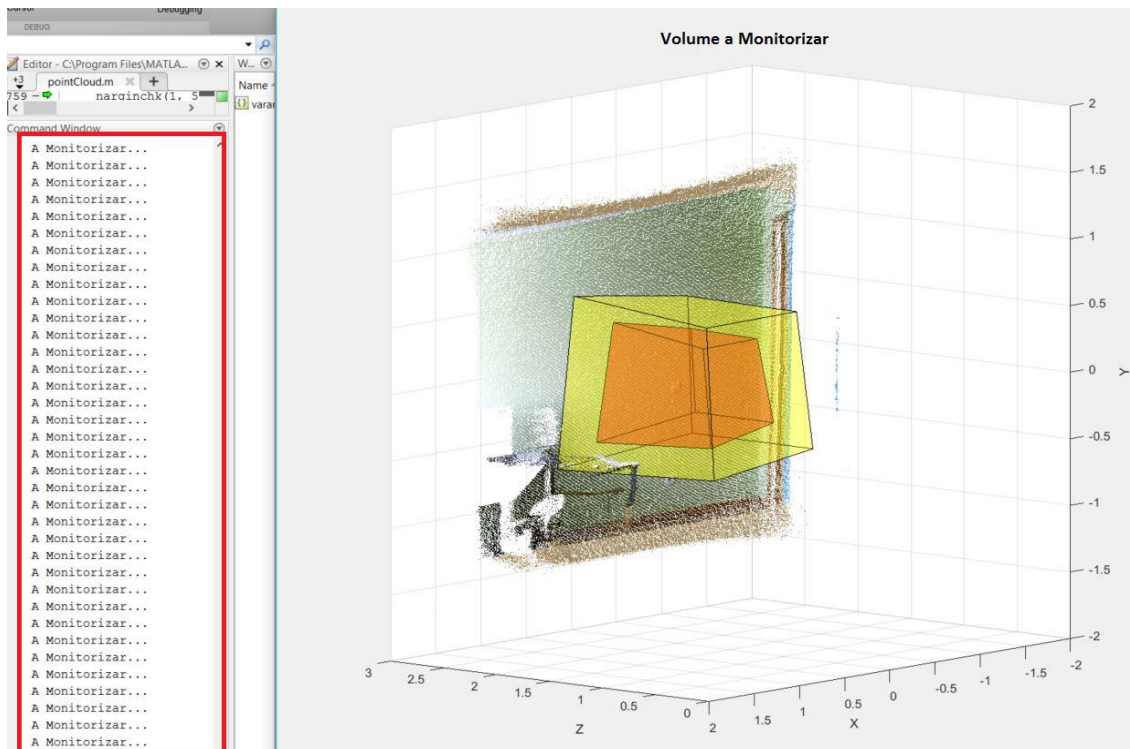


Figura 52. Funcionamento da fase de monitorização (ciclo em curso).

Como se verifica na Figura 52, dentro do retângulo a vermelho (do lado esquerdo) está presente a informação da consola do programa, que informa o utilizador que o programa se encontra em monitorização.

Na Figura 53 é apresentado um exemplo da zona de perigo (a amarelo) a ser infringida.

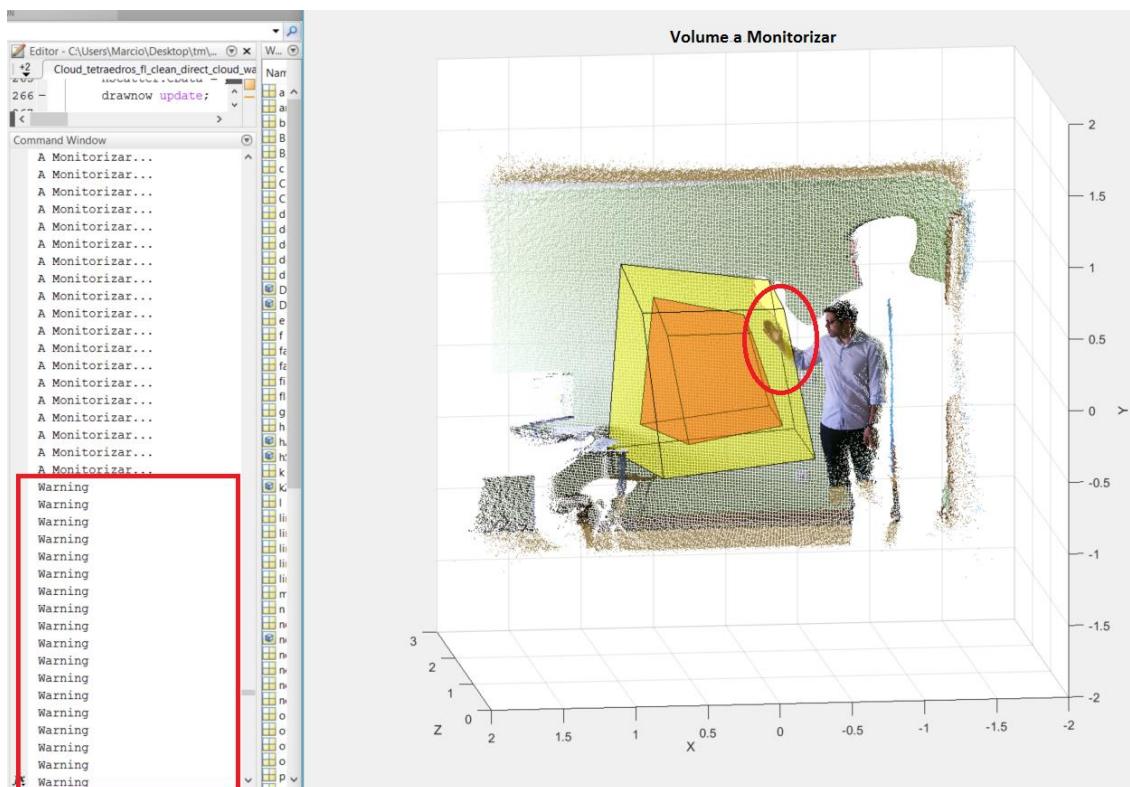


Figura 53. Funcionamento da fase de monitorização (zona de perigo infringida).

Através da Figura 53, visualiza-se no retângulo vermelho, a informação “Warning”, que informa o utilizador que a zona de perigo foi infringida. Dentro do círculo vermelho pode-se visualizar a mão dentro da zona amarela, o que indica que a zona de perigo foi infringida.

A Figura 54 representa um exemplo de uma intrusão na zona interdita, o que resultará numa mensagem de aviso ao utilizador.

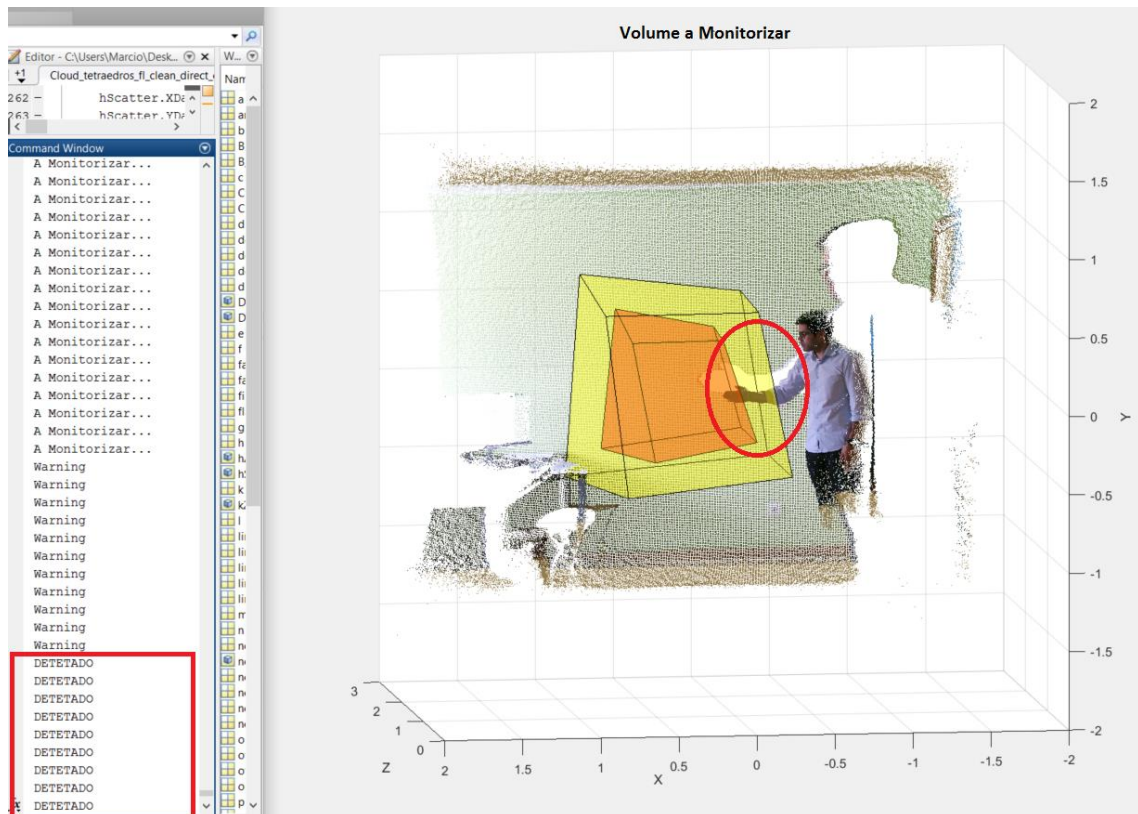


Figura 54. Funcionamento da fase de monitorização (zona interdita infringida).

Na Figura 54 é possível visualizar, dentro do círculo vermelho, uma intrusão na zona interdita. Consequentemente, surge do lado esquerdo da imagem uma mensagem apresentada na consola do programa a informar o utilizador que existe uma intrusão.

De forma a testar a aplicação relativamente à evasão do interior da zona interdita, colocou-se um objeto dentro da zona interdita e, posteriormente, colocou-se a aplicação em funcionamento. Pode-se visualizar o comportamento da aplicação através das Figuras 55 e 56.

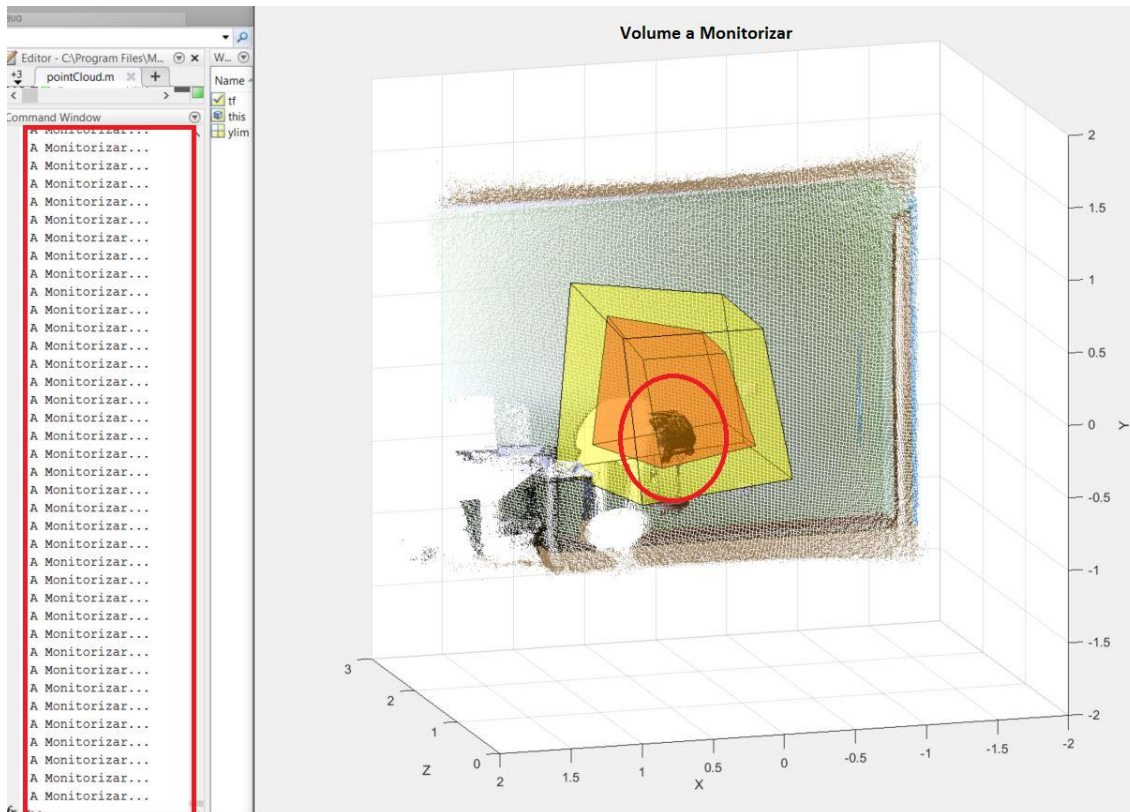


Figura 55. Funcionamento da fase de monitorização (objeto dentro da zona interdita).

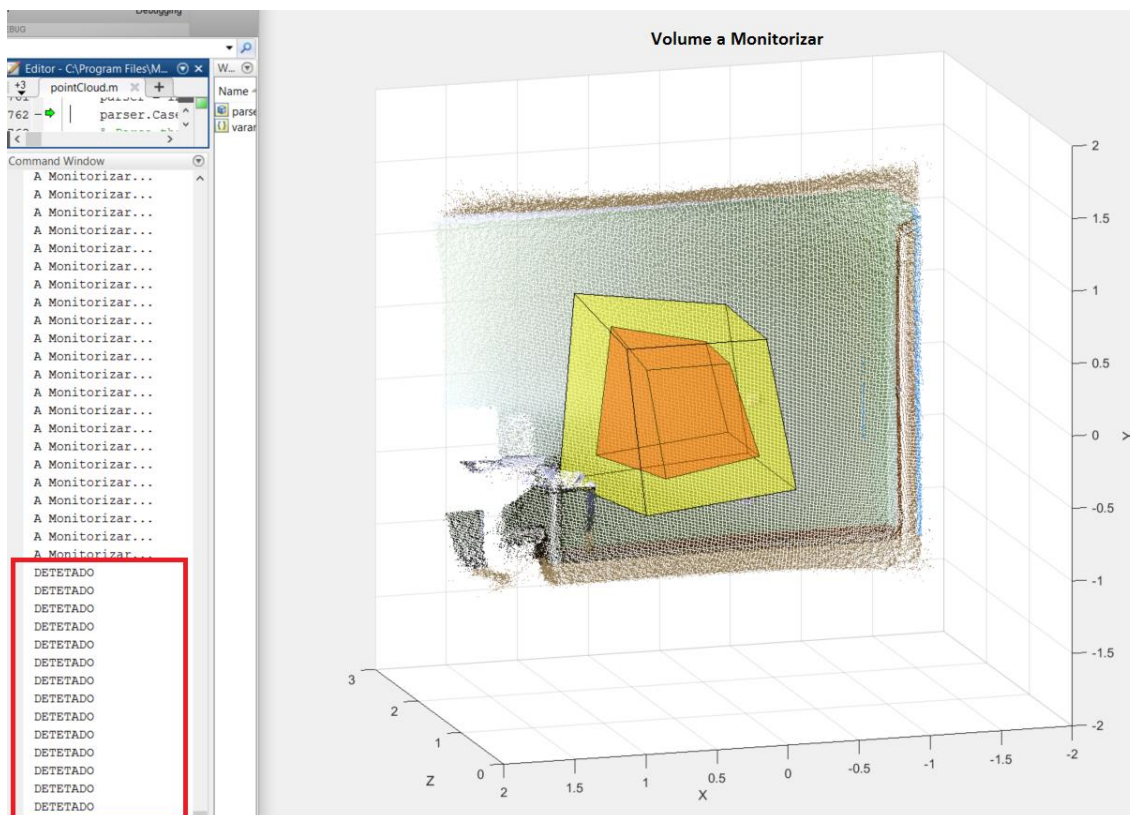


Figura 56. Funcionamento da fase de monitorização (objeto retirado da zona interdita).

Para os testes ilustrados nas Figuras 55 e 56, foi inserido um objeto na zona de volume interdito, antes de iniciar a aplicação no modo de monitorização, estando o objecto presente na fase de calibração. O passo seguinte foi iniciar a aplicação em modo de monitorização, como observado na Figura 55. O objeto foi retirado do interior da zona interdita e, tratando-se de uma evasão do interior da zona protegida, foi exibida uma mensagem na consola do programa, criando um alerta para o utilizador.

A monitorização descrita ao longo desta secção é exclusivamente direccionada para objetos que se encontram estáticos, no interior da zona interdita. De forma a que o sistema possa ser implementado em ambiente dinâmico, ou seja, quando há objectos em movimento na zona interdita e que são parte integrante do ambiente inserido no volume a monitorizar, efectuaram-se algumas alterações ao nível de *software*, bem como na representação da área a ser monitorizada, tal como se pode observar na Figura 57.

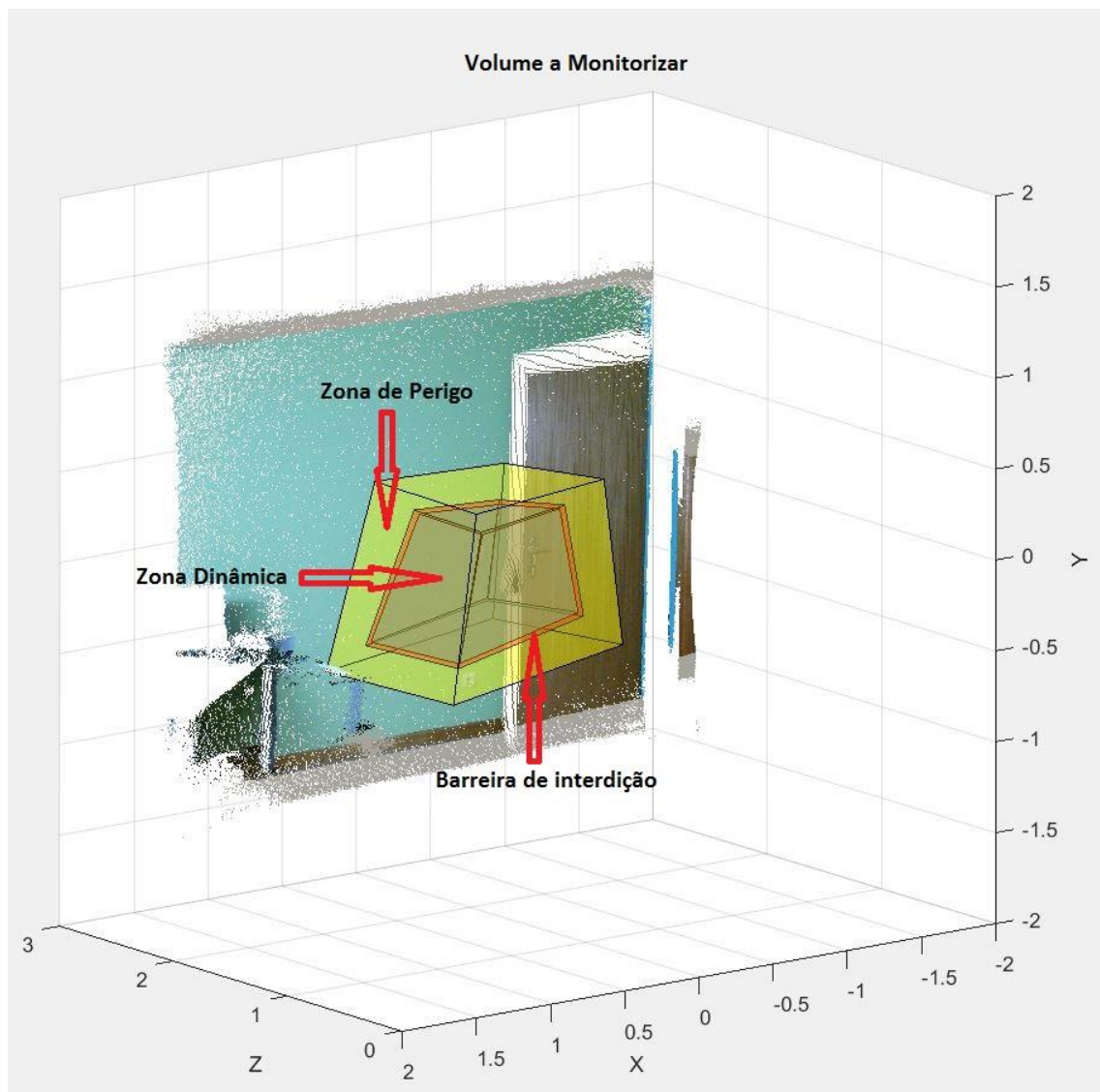


Figura 57. Funcionamento da fase de monitorização (zona dinâmica).

Na Figura 57 está ilustrada a amarelo a zona de perigo, o contorno vermelho que indica as barreiras exteriores à zona interdita, e por fim, a envolvente representada a cinzento que diz respeito à zona de ocupação dinâmica.

Na Figura 58, está representada uma intrusão na zona de perigo e na zona interdita.

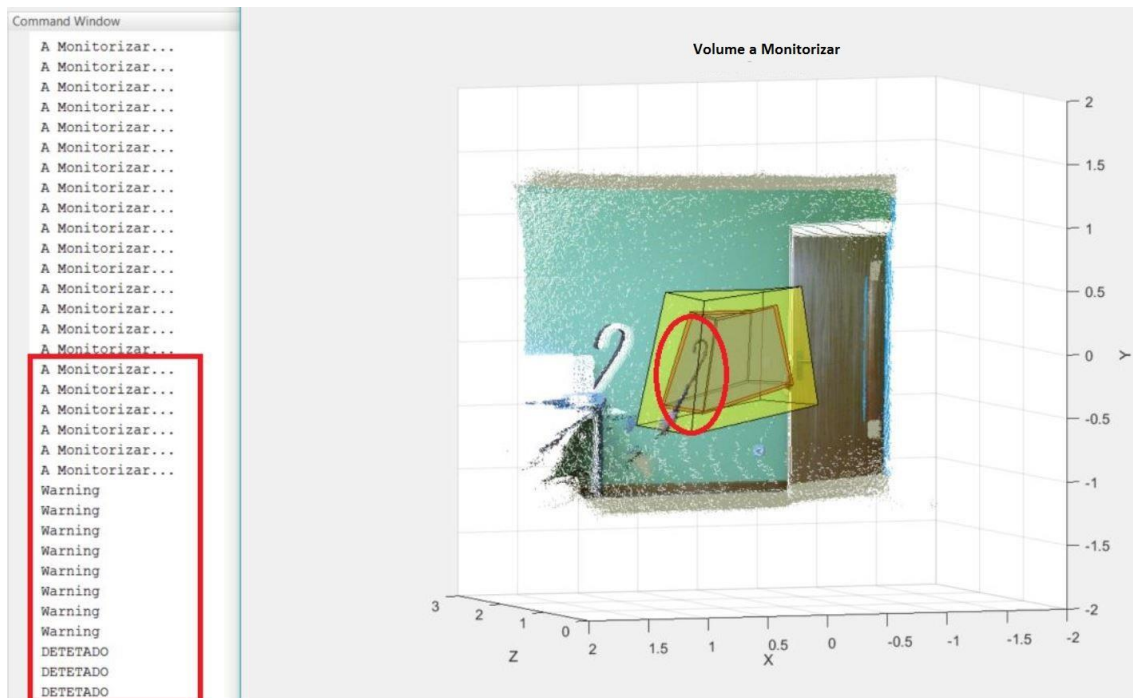


Figura 58. Funcionamento da fase de monitorização (intrusão na zona dinâmica).

Através da Figura 58, pode-se verificar que existe uma intrusão na zona interdita, sendo esta detetada e indicada pela mensagem na consola do programa. À semelhança da deteção em ambiente estático, a deteção em ambiente dinâmico dispõe de uma zona de perigo e de uma zona de interdição, nas quais pode ocorrer uma intrusão ou evasão. A diferença mais significativa entre o método de deteção para ambiente estático e o método para ambiente dinâmico, é o facto de no método para ambiente dinâmico o algoritmo de deteção considerar mais determinante para a decisão de intrusão/evasão a violação das barreiras exteriores. Assim, no método para ambiente dinâmico, a função principal é a monitorização da violação das barreiras de interdição, não sendo valorizadas da mesma forma as variações que ocorrem na ocupação do volume interior.

4.4 Fluxograma geral do *software*

O fluxograma geral do *software* representado na Figura 59 ilustra todo o funcionamento da aplicação de monitorização desenvolvida. A divisão deste fluxograma pode ser realizada em três fases principais. A fase inicial em que se efetua a calibração, a fase referente à monitorização contínua e, por fim, a concretização do objetivo principal que se refere à fase da deteção efectiva. Cada uma destas fases não pode funcionar individualmente, uma vez que possuem dependências que exigem que sejam sequenciais. Inseridas nestas três fases mais gerais, encontram-se ainda outras acções específicas importantes, que estão associadas à aquisição de dados da câmara e ao método de divisão do volume em tetraedros (Triangulação Delaunay 3D).

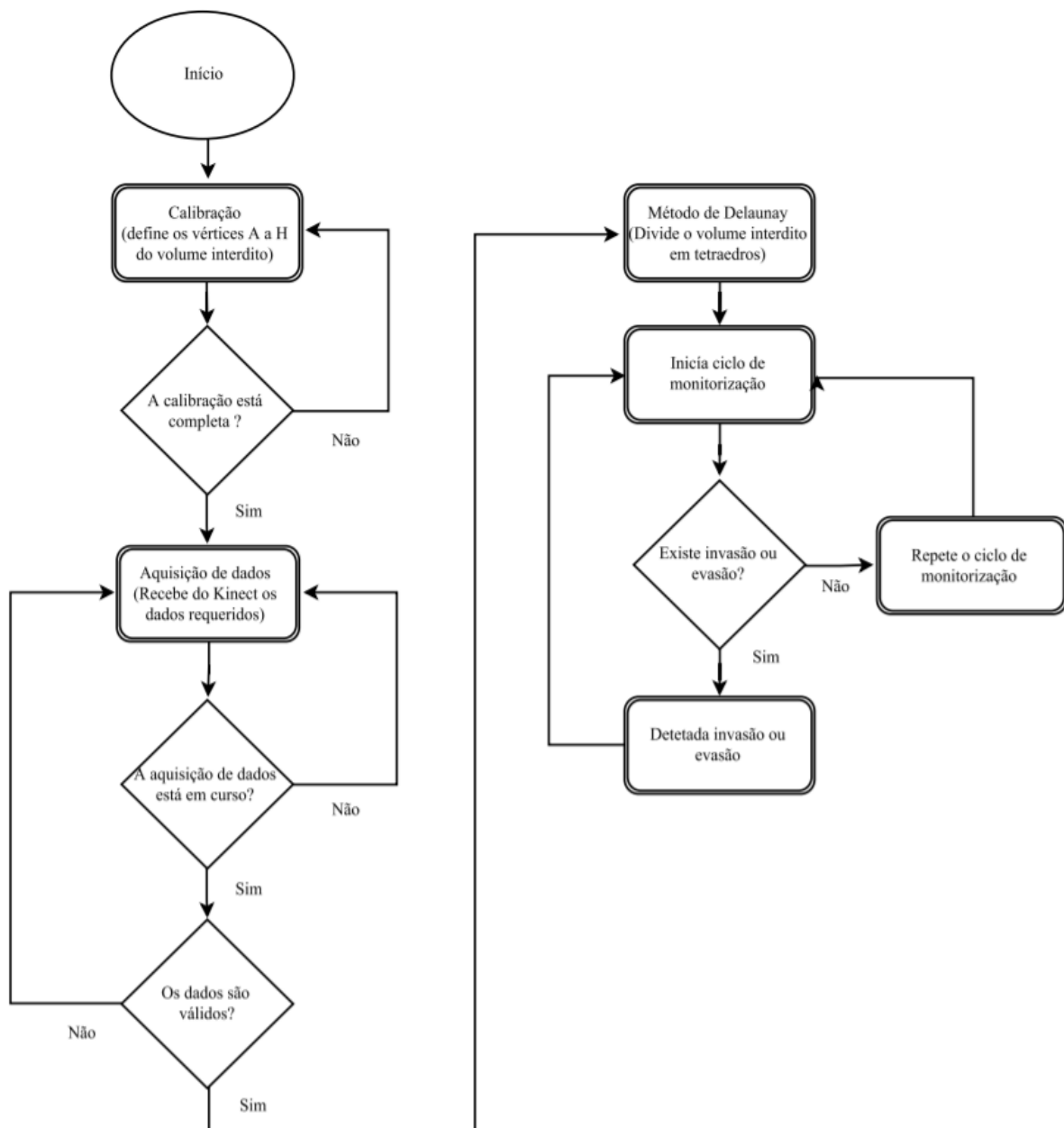


Figura 59. Fluxograma geral do programa.

4.5 Performance temporal do *software* e *hardware*

Existem inúmeros aspetos a ter em conta na obtenção de uma performance temporal aceitável, que resulte num bom funcionamento da aplicação de monitorização. Esta performance temporal resulta, maioritariamente, do modo como o programa se encontra estruturado, da velocidade de envio de informação por parte do sistema Kinect V2 e da capacidade de processamento do computador onde está a ser executada a aplicação. A performance temporal vai determinar as capacidades de resposta de um sistema de monitorização em ambiente industrial, por exemplo no tempo máximo de resposta na acção de parar uma máquina após a deteção de uma intrusão no volume interdito.

4.5.1 Resultados temporais da aquisição de dados

Os resultados relativos à aquisição de dados são obtidos no ciclo de monitorização e dizem respeito ao tempo que decorre entre o pedido de dados por parte do *software* e a respetiva resposta por parte do *hardware*. Pode-se observar o resultado deste teste na Figura 60.

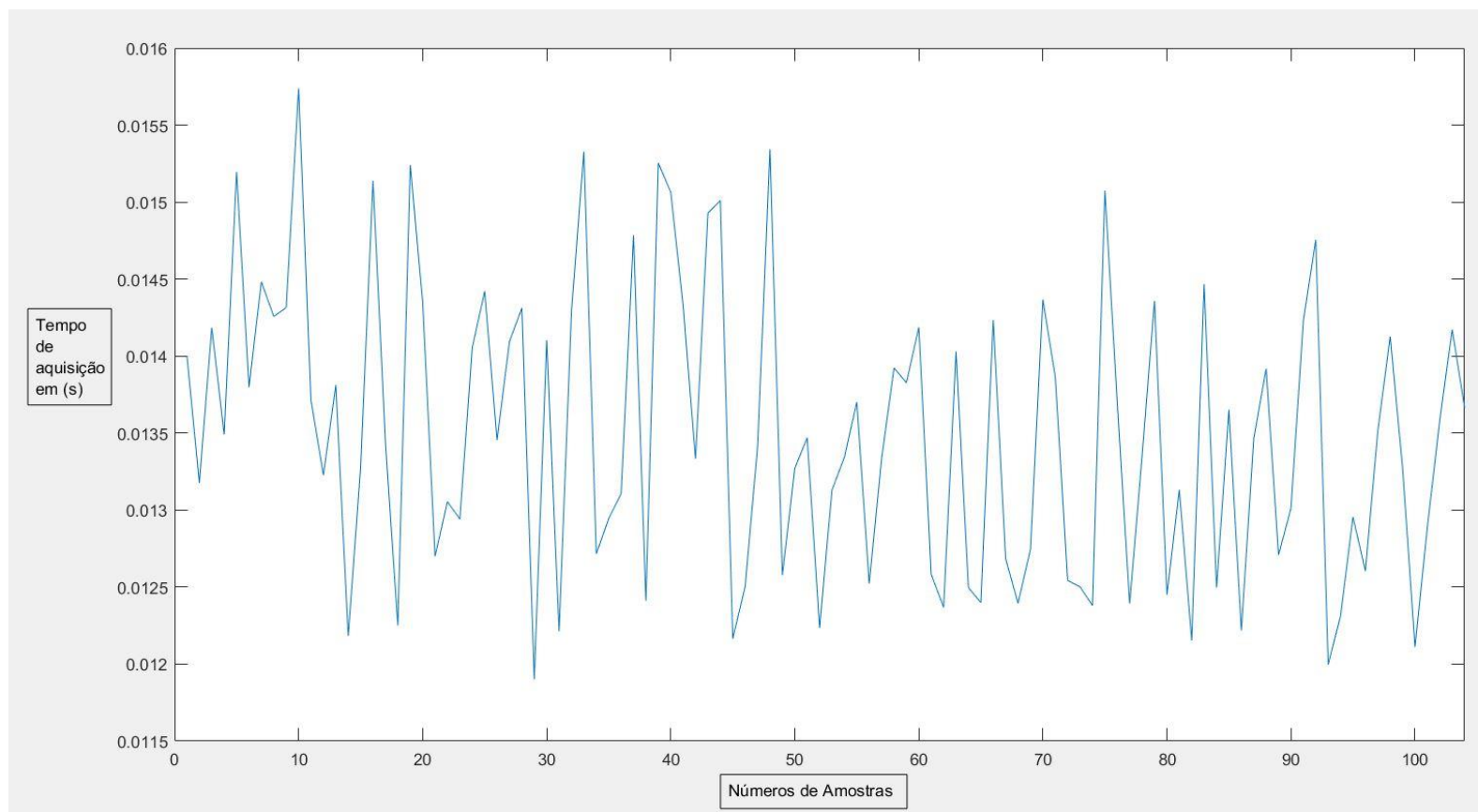


Figura 60. Tempo necessário para que o Kinect V2 responda a uma solicitação de dados.

Por observação da Figura 60, conclui-se que o tempo que o Kinect V2 necessita para responder a uma solicitação de dados varia entre cerca de 11,9 ms e 15,8 ms, sendo este tempo considerado ótimo para os objectivos da aplicação desenvolvida.

4.5.2 Resultados temporais do ciclo de monitorização

A duração do ciclo de monitorização é o fator que determina o tempo máximo entre uma possível detecção e a próxima detecção, ou seja, é o tempo máximo que o sistema vai necessitar para monitorizar o volume interdito e verificar se existiu ou não uma intrusão/evasão. A Figura 61 apresenta o comportamento do *software* desenvolvido em termos do tempo necessário para completar um ciclo de monitorização e possível detecção.

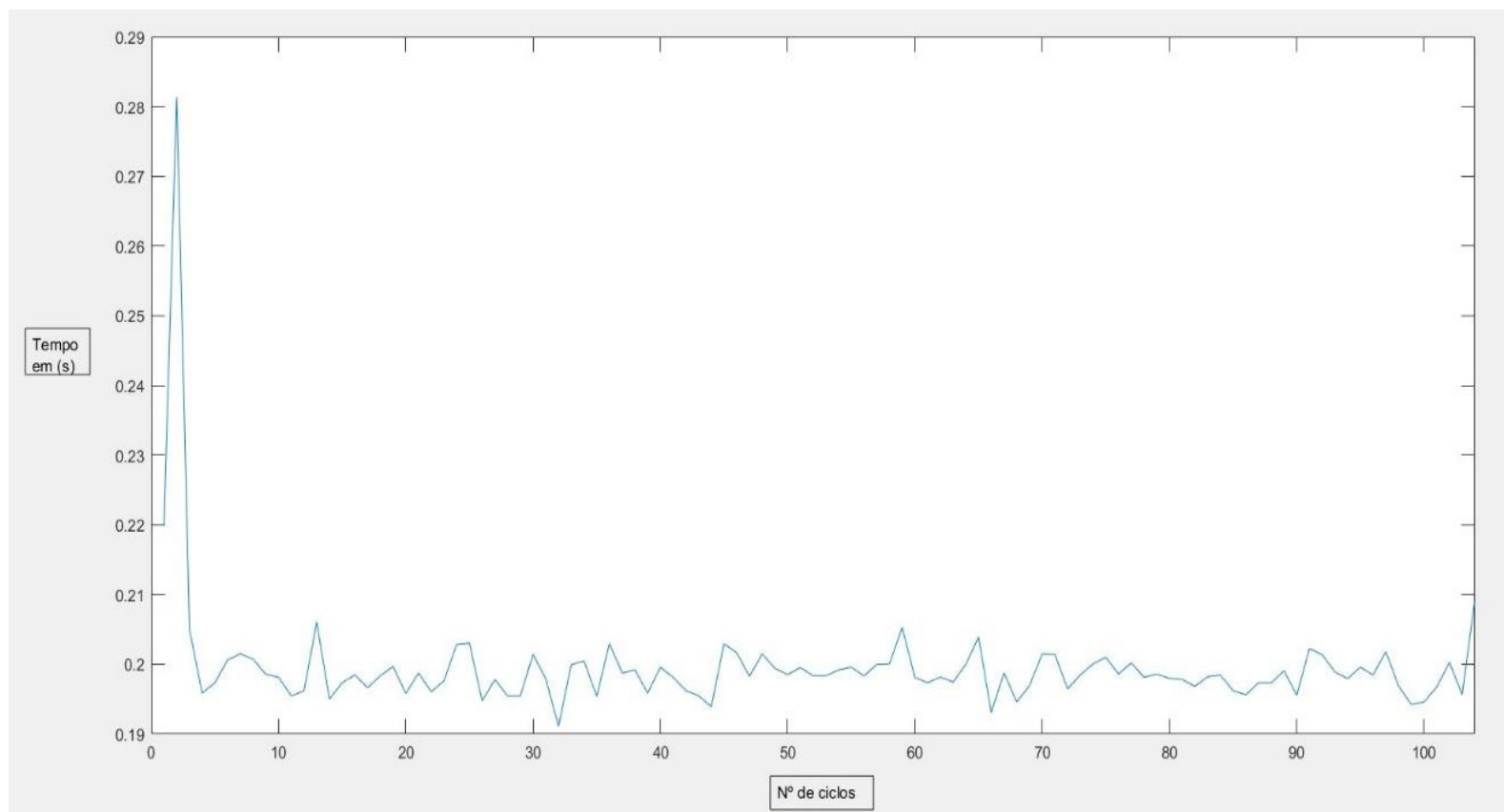


Figura 61. Tempo necessário para a aplicação executar um ciclo de monitorização completo.

Observando a Figura 61, verifica-se que o primeiro ciclo é mais demorado, durando cerca de 280 milissegundos. Este facto deve-se à declaração inicial de variáveis dentro do ciclo. Os restantes ciclos exibem uma duração que varia entre 191 milissegundos e 209 milissegundos. Estes tempos incluem a atualização contínua da interface gráfica. Concentrando todo o processamento na deteção, é possível reduzir este tempo para menos de 100 milissegundos.

Com base nestes valores, considera-se que a aplicação pode ser utilizada para uma monitorização contínua em tempo real. O tempo de resposta máximo a uma intrusão será nesta implementação de cerca de 100 milissegundos.

4.5.3 Performance do computador utilizado

Para enquadrar os resultados anteriores, é necessário especificar o *hardware* utilizado, uma vez que o tempo de monitorização está diretamente associado à capacidade de processamento. Desta forma, apresentam-se as especificações do computador utilizado nas experiências anteriores:

Processador: Intel(R) Core(TM) i5-4210h CPU @ 2.90Ghz, 2 núcleos, 4 processadores

Ram: 8,00 GB

Tal como referido anteriormente, a capacidade de processamento está diretamente associada ao tempo de execução do ciclo da aplicação. Assim sendo, é possível melhorar, caso seja necessário, os resultados obtidos anteriormente, com a implementação da aplicação num processador com maior velocidade de execução. Adicionalmente, os tempos indicados foram obtidos utilizando o Matlab com *software* não otimizado e em modo interpretado. Estes tempos podem ser significativamente reduzidos se o algoritmo de deteção for otimizado em termos de velocidade e executado num formato pré-compilado. O processo de aquisição de dados por parte do Kinect não representa a limitação principal.

4.6 Implementação em ambiente industrial

De modo a integrar a aplicação num ambiente real, procedeu-se à sua instalação e calibração num ambiente que pretende representar o ambiente industrial, mais especificamente numa plataforma onde se encontra inserido um braço robótico (Figura 62). Foi criado um *script* de execução para o braço robótico, onde este percorre o interior do espaço interdito na maioria do seu volume, tendo os testes sido realizados com o robô a executar esta rotina programada.

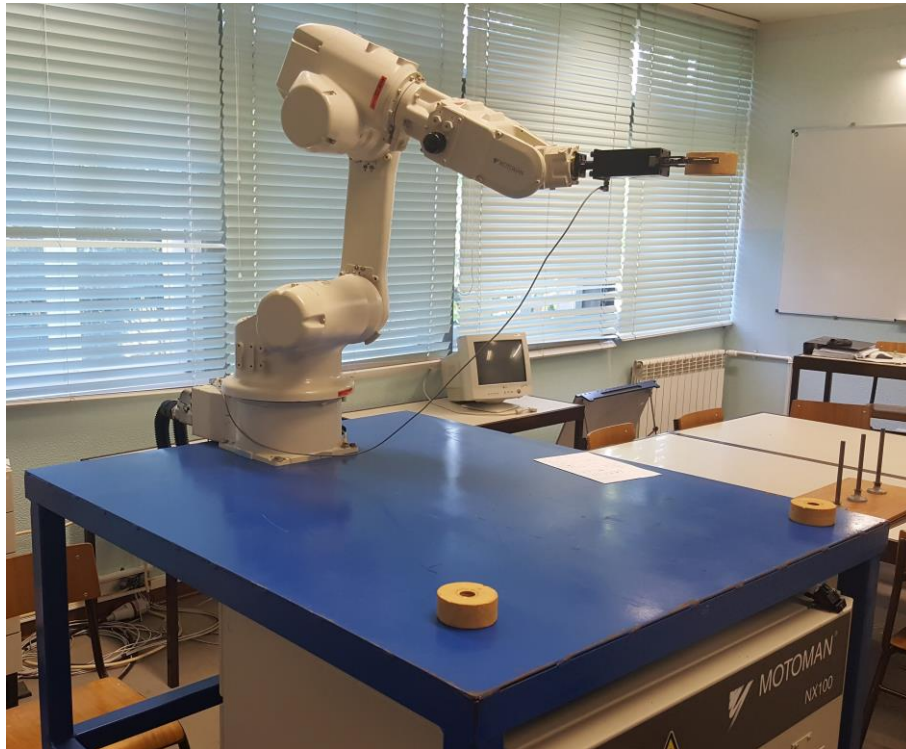


Figura 62. Robô Motoman usado para a monitorização.

Como forma de testar uma ação de monitorização específica, criou-se um volume interdito em torno da plataforma onde o braço robótico se encontra instalado (Figura 63). Dado que esta plataforma possui uma base opaca de metal, não se considerou incluir na zona interdita o volume abaixo da plataforma de metal, uma vez que essa região não seria visível pela câmara, que se colocou acima da referida base (Figura 63). A zona interdita definida acima da base de metal pode ser observada na Figura 63.

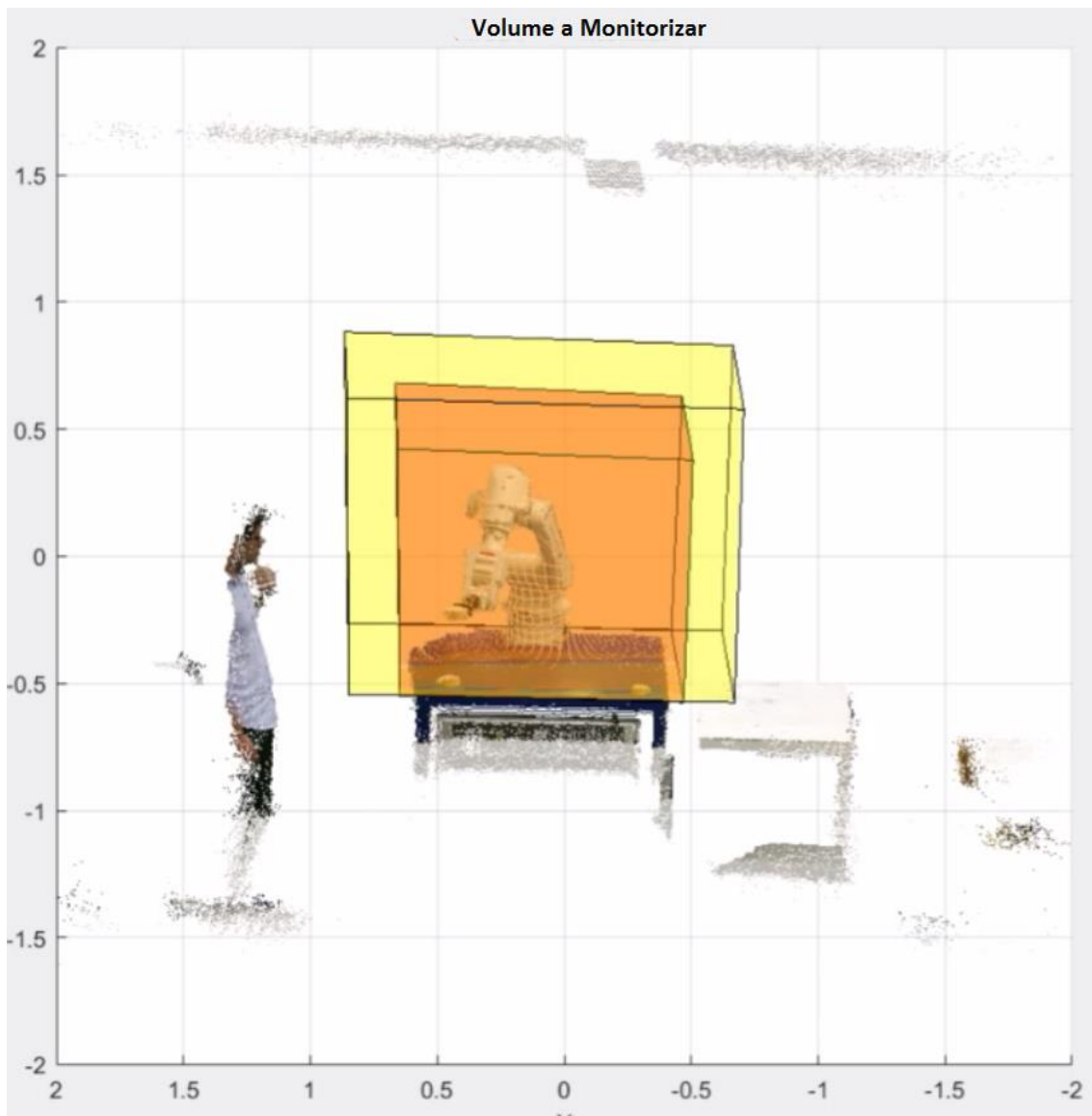


Figura 63. Volume interdito definido.

Como se verifica na Figura 63, a região exterior (a amarelo) é a parte que diz respeito à aproximação do espaço interdito, designada de zona de perigo. A definição deste conceito serve para que seja possível avisar o utilizador que se está a aproximar do espaço interdito. No interior da zona de perigo situa-se o volume interdito (volume a vermelho). Uma vez entrando no espaço interdito, o utilizador está a colocar em risco a sua segurança ou a segurança do processo e/ou dos equipamentos.

4.6.1 Resultados da implementação em ambiente industrial

Com a finalidade de testar a aplicação no ambiente de testes, foi-se infringindo as duas zonas (zona de perigo e zona interdita). Nas Figuras 64, 65 e 66 constata-se a eficácia da aplicação na monitorização e deteção de intrusões.

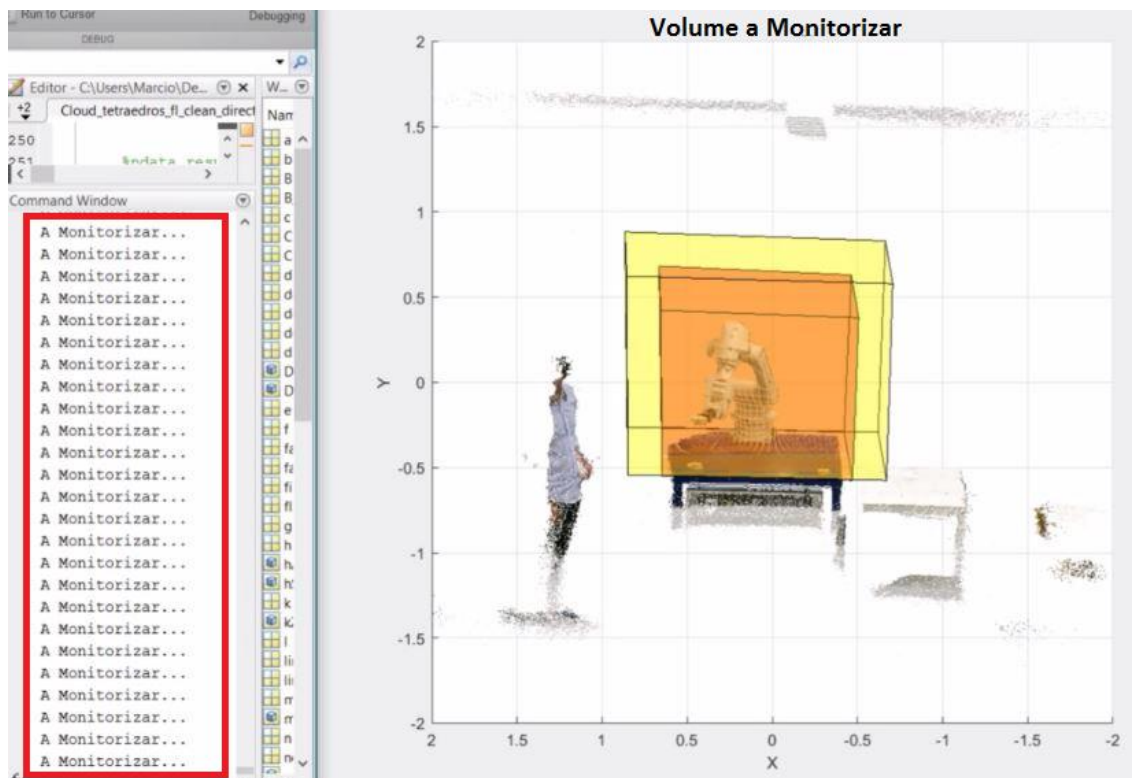


Figura 64. Monitorização em curso.

Observando a Figura 64, identificam-se as zonas de perigo e interdita, que não estão a ser infringidas. Daí resulta na consola da aplicação uma mensagem que informa o utilizador com a seguinte mensagem: “A Monitorizar”.

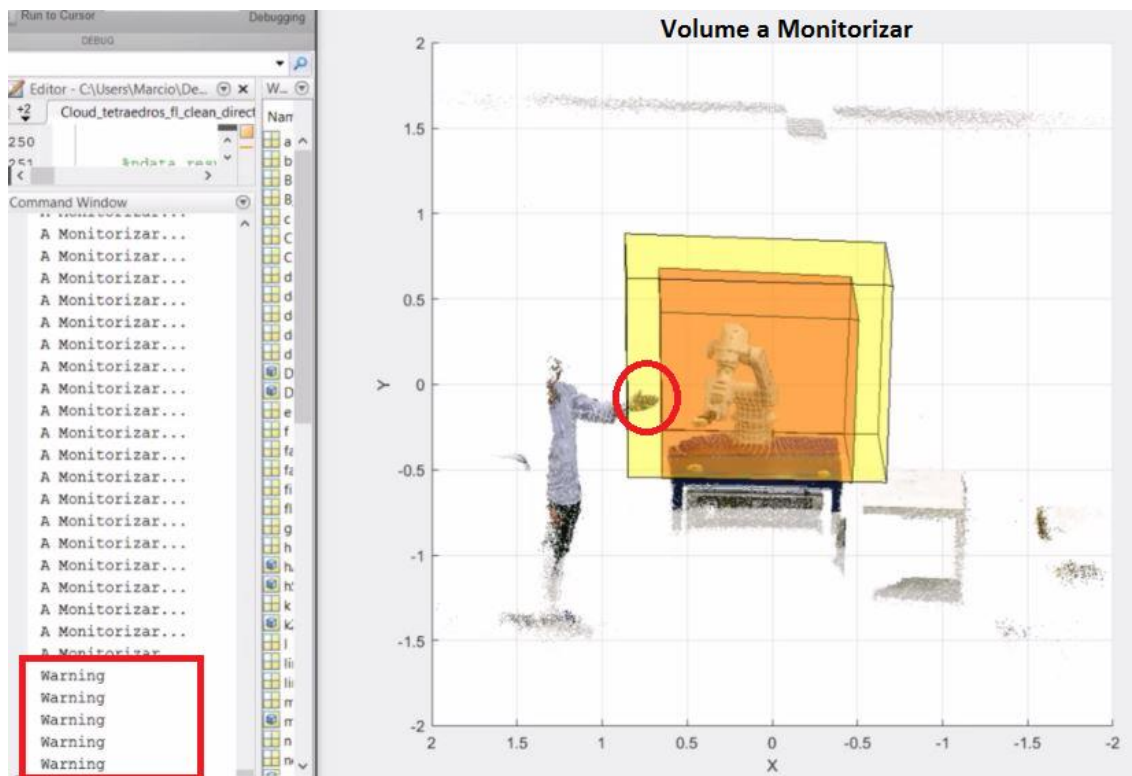


Figura 65. Volume de perigo infringido.

Na Figura 65, pode-se observar que o utilizador está a infringir a zona de perigo e, na consola do programa (lado esquerdo), surge a mensagem “Warning”. Juntamente com este aviso visual, surge um sinal sonoro e uma luz de sinalização amarela, que indica o estado de perigo.

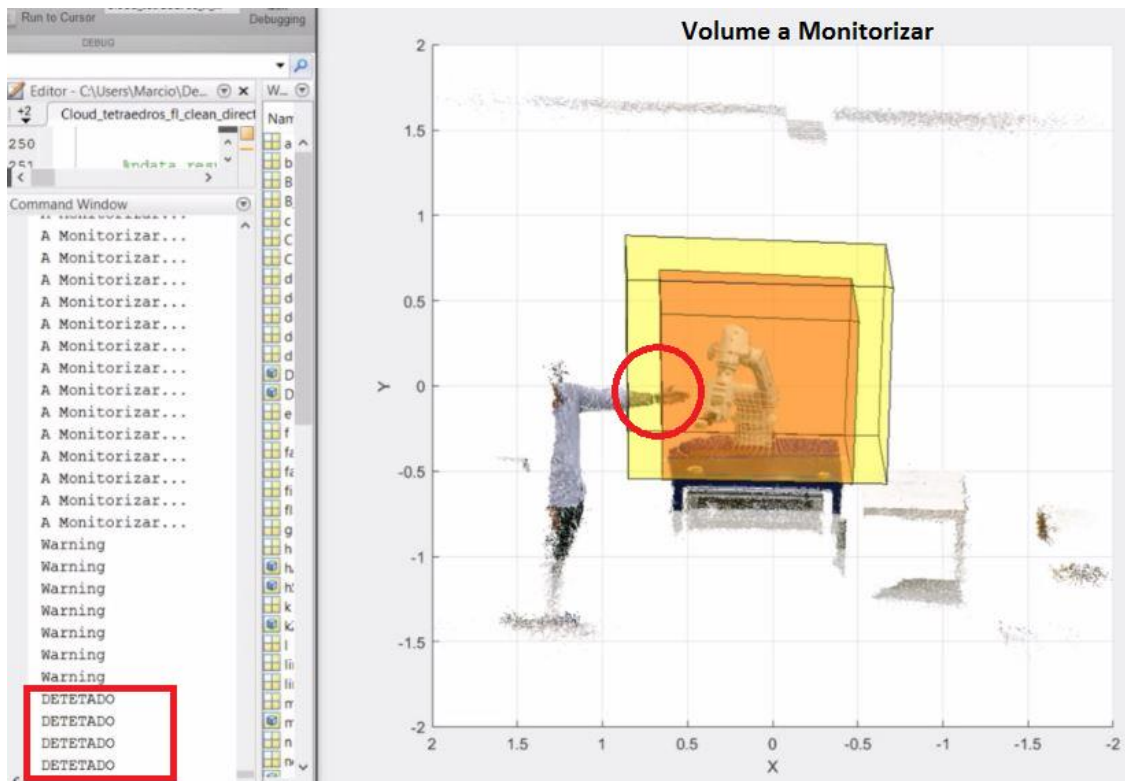


Figura 66. Volume interdito infringido.

Analisando a Figura 66, observa-se que a zona interdita foi infringida, sendo gerada a mensagem “DETETADO” (lado esquerdo da imagem). Com este aviso, surge também um sinal sonoro com uma frequência diferente do sinal emitido no estado de “Warning” e a sinalização apresentada por uma luz vermelha. Existe ainda um relatório que é gerado com a data e hora de quando ocorreram as intrusões (Figura 67).

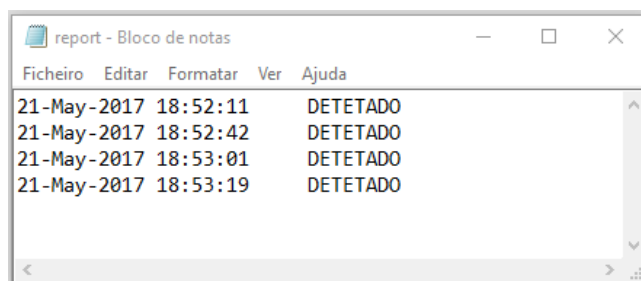


Figura 67. Relatório gerado quando há intrusão/evasão no/do espaço interdito.

4.6.2 Condições dos resultados obtidos na implementação em ambiente industrial

Os resultados da implementação em ambiente industrial foram registados com o braço robótico estático numa posição fixa e, posteriormente, com o mesmo em funcionamento executando o percurso e acções definidas no *script* pré-programado.

Foi também testado com sucesso o funcionamento do sensor Kinect V2 com luz natural e luz artificial na sala onde o braço robótico se encontra instalado

4.7 Interface com o utilizador

Com o objetivo de facilitar a interação do utilizador com o programa em funcionamento, bem como simplificar a execução das opções mais relevantes, foi criada uma interface gráfica, de acordo com as várias funcionalidades que o programa permite executar.

Deste modo, é possível seleccionar as opções “modo estático” ou “modo dinâmico” (Figura 68) e, seleccionado o modo, é possível colocar o programa em modo de monitorização.

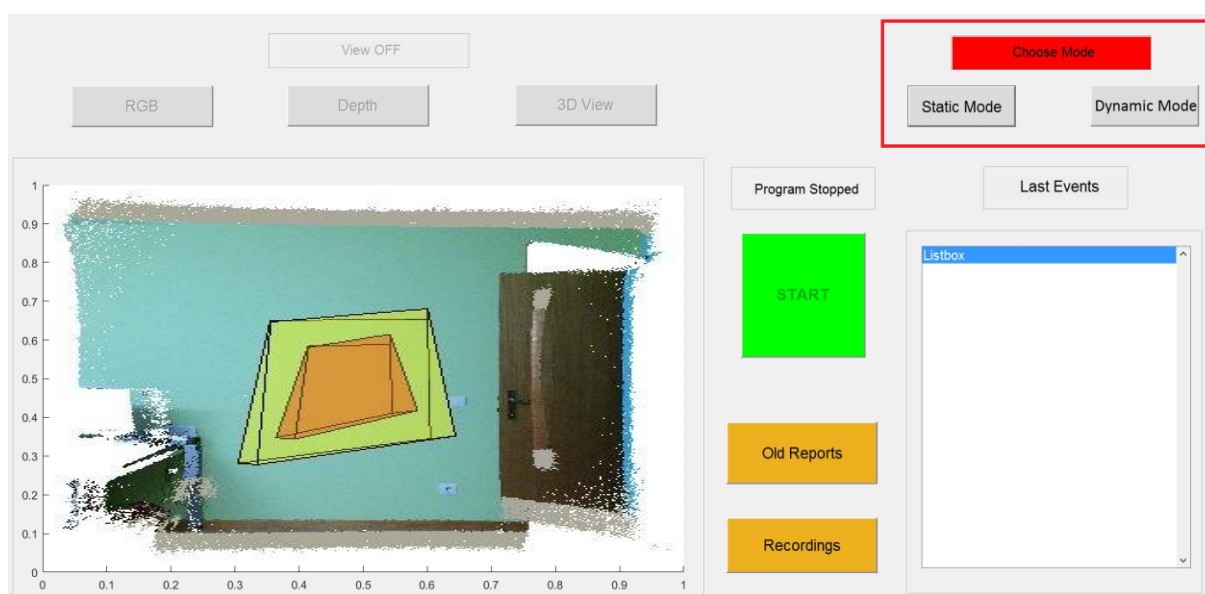


Figura 68. Seleção do modo de funcionamento.

Após a seleção do modo de funcionamento pretendido, o botão de “START” fica disponível para iniciar o programa, permitindo ao utilizador dar início ao processo de monitorização contínua, de acordo com o modo escolhido previamente (Figura 69).

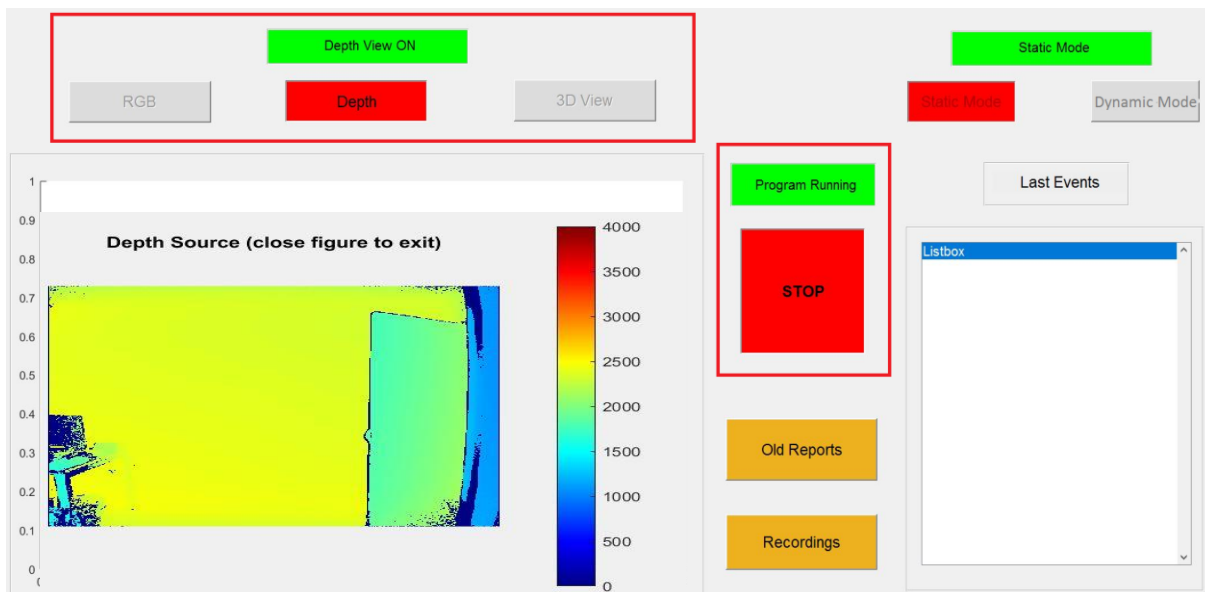


Figura 69. Iniciar o modo de monitorização.

Com o programa em modo de monitorização, é possível seleccionar o modo de visualização pretendido pelo utilizador, tendo à disposição o modo “RGB” que permite visualizar a imagem a cores diretamente da câmara, o modo “Depth” no qual se pode visualizar a imagem 2D e a respetiva profundidade e, por fim, o modo “3D View” que apresenta a imagem 3D adquirida e representada por uma nuvem de pontos.

Optou-se ainda por adicionar dois botões que permitem visualizar relatórios de ocorrências anteriores, bem como as respetivas gravações (Figura 70).

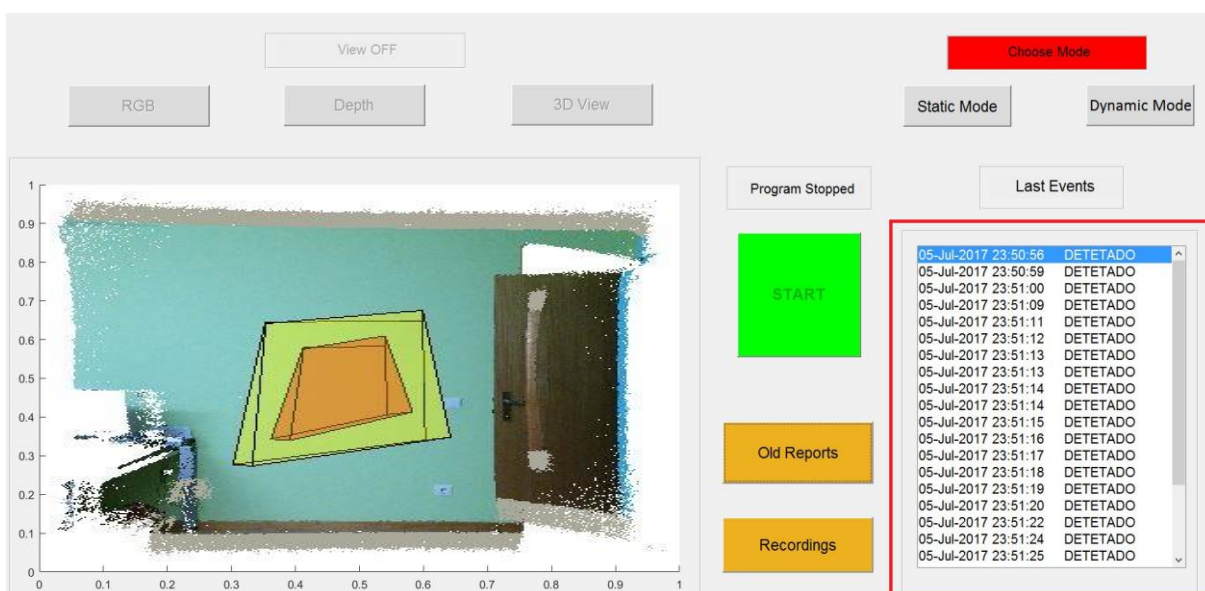


Figura 70. Relatório de ocorrências e gravações.

O botão “Old Reports” permite visualizar relatórios de ocorrências anteriores, enquanto o botão “Recordings” permite obter as respectivas gravações.

4.8 Implementação de alarme

Para demonstrar como pode ser utilizada a informação de intrusão ou evasão do espaço a ser monitorizado numa aplicação industrial concreta, foi implementado um sistema simples que integra um alarme visual e um alarme sonoro. Este sistema consiste numa placa Arduino Uno [66] diretamente ligada ao Matlab, através de uma interface USB. A interface criada é baseada num pacote de de *software* de suporte do Matlab, criado especificamente para a interligação com os sistemas Arduino [67].

4.8.1 Método de alarme

O alarme implementado consiste num sinalizador luminoso e outro sonoro, permitindo assim que o utilizador facilmente seja informado de uma intrusão ou evasão no/do espaço monitorizado (Figura 71) e do seu tipo. Mais especificamente, quando há uma intrusão ou evasão do espaço, o sinal luminoso e o sinal sonoro são ativados. O sinal luminoso é amarelo para uma intrusão na zona de perigo e vermelho para uma intrusão na zona interdita. Por seu lado, o sinal sonoro possui uma frequência para uma intrusão na zona de perigo e outra frequência quando existe intrusão na zona interdita.

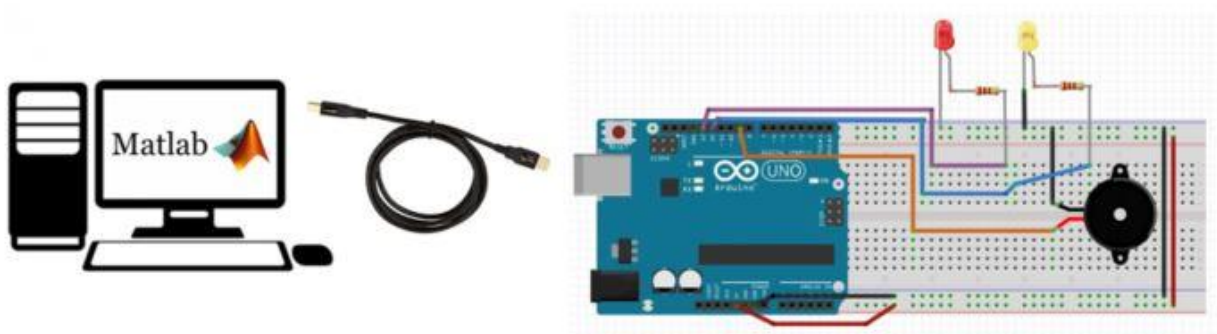


Figura 71. Ilustração da ligação do arduino ao Matlab.

O funcionamento do sistema de alarme em termos de *software* pode ser explicado através do fluxograma apresentado na Figura 72.

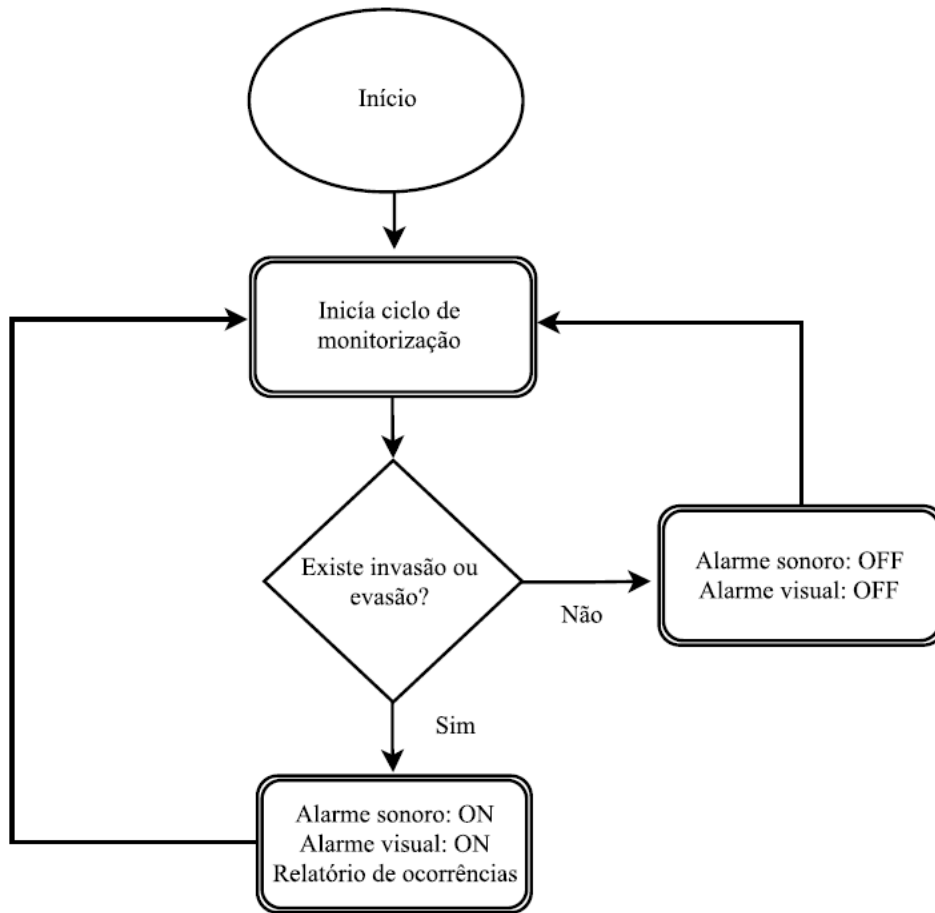


Figura 72. Fluxograma do sistema de alarme.

5 Conclusão e trabalhos futuros

5.1 Conclusão

Este trabalho teve por objetivo a monitorização de áreas e volumes em tempo real, com a utilização de uma câmara RGB-D, mais especificamente o sistema Kinect V2. Inicialmente elaborou-se um estudo e levantamento da informação relevante e da legislação em vigor, que enquadra a implementação e utilização de sistemas de monitorização com base em vídeo, bem como sobre as tecnologias e sistemas existentes atualmente no mercado com esse objetivo.

Estudaram-se as técnicas de aquisição de imagem 3D e, após uma comparação entre os parâmetros mais específicos das várias técnicas, optou-se pela tecnologia mais inovadora e com maior rigor na medição de profundidade, a tecnologia TOF. Após a escolha da tecnologia, foi efetuada uma seleção cuidada do *hardware* disponível no mercado, de forma a satisfazer as necessidades de integração no projeto.

Com vista ao desenvolvimento da aplicação de monitorização, o Matlab foi o ambiente de programação escolhido. Em conjunto com o SDK disponibilizado pela Microsoft e uma interface específica criada em linguagem C, que permite a transferência de dados entre o sistema Kinect V2 e computador onde se encontra instalado o Matlab, foi possível avançar para a fase de implementação de uma aplicação concreta.

Após a integração com sucesso do Kinect V2 no ambiente de programação Matlab, e com a confirmação da correta aquisição e disponibilização das imagens adquiridas, começou-se por criar uma aplicação de calibração, onde se debateu a forma mais simples e rápida de fazer a representação da zona interdita a monitorizar. Concluiu-se que uma forma simples e rápida seria o utilizador posicionar-se em frente da câmara e utilizar apenas os seus dedos para definir os pontos que constituem os vértices do volume interdito a monitorizar.

Relativamente ao desenvolvimento da fase de monitorização, esta foi a parte do projeto na qual houve necessidade de implementar um algoritmo mais elaborado, uma vez que é necessário receber toda a informação proveniente da câmara, analisar essa informação em tempo real e implementar as ações que permitem à aplicação efetuar a monitorização da zona 3D, previamente definida pelo utilizador na aplicação de calibração.

A fim de permitir uma monitorização intuitiva para o utilizador, e reproduzindo a função do sistema a funcionar em ambiente industrial, foi criado um sistema que gera um alarme visual e sonoro, sempre que a zona de perigo ou a zona interdita são invadidas. É ainda produzido, em ficheiro, um relatório de ocorrências, que indica o dia e a hora exata das ocorrências registadas. A ativação de alarmes pode ainda incluir ações de parar máquinas individuais e linhas de produção completas.

Finalmente, simulando um ambiente industrial, efetuaram-se testes ao sistema no volume de operação um braço robótico em funcionamento contínuo. O sistema desenvolvido cumpriu todas as especificações pré-definidas, em cenários estáticos e dinâmicos, podendo assim ser usado com sucesso em vários cenários de monitorização 3D, incluindo como dispositivo de proteção industrial baseado em visão.

5.2 Proposta de desenvolvimento futuro

Sendo possível avançar para muitos outros tipos de aplicações com um sistema Kinect V2 e ferramentas de *software* consolidadas, uma possibilidade futura não passa apenas pelo *software*, mas também por alterações de base ao nível de *hardware*. Mais propriamente, propõe-se a instalação de mais sistemas Kinect V2 em torno da região a monitorizar (Figura 73), satisfazendo, deste modo, a necessidade de mapear por completo o respetivo volume. Esta necessidade pode surgir pela limitação de com um sensor apenas poderem existir zonas de sombra, cujas possibilidades de mapeamento de profundidades se encontram limitadas pela oclusão sofrida por partes dos objetos mais próximas da câmara (única).

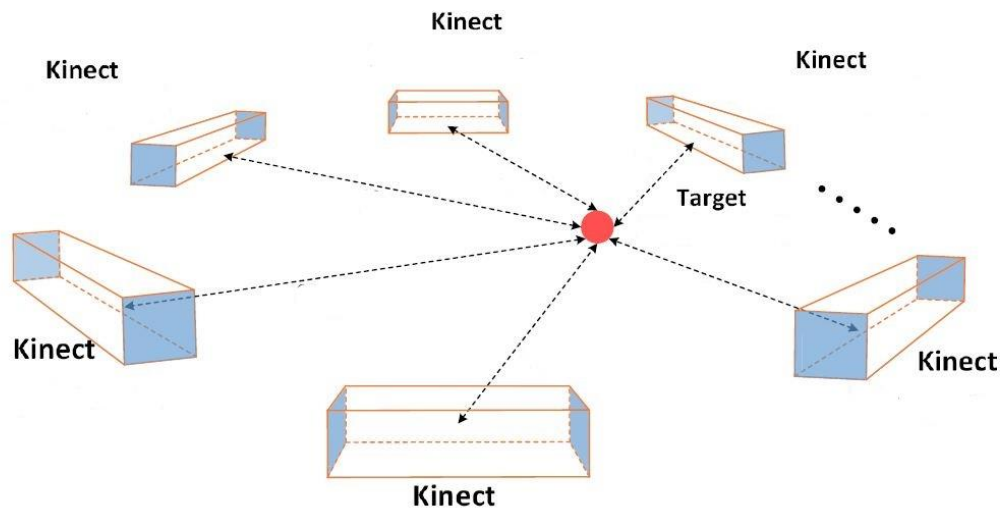


Figura 73. Monitorização utilizando vários sistemas Kinect V2, adaptado de [68].

A monitorização a ser efetuada deste modo, necessita de cuidados específicos a ter em conta, nomeadamente, a localização relativa entre os vários sensores, bem como a sua localização individual relativamente à região alvo. Para além disso, com o sistema devidamente instalado, com a ajuda de métodos de fusão de informação, a exatidão e precisão da localização dos pontos a monitorizar podem ser de melhor qualidade. Este é um trabalho que irá naturalmente consumir mais recursos, tanto em *hardware* como em *software*, e a sua interface poderá ter que ser construída especialmente para esta aplicação, dada a quantidade de portas USB 3.0 necessárias à sua implementação.

Referências Bibliográficas

- [1] ISO 3864-1 Símbolos gráficos. Cores de segurança e sinais de segurança. Parte 1: Critérios de desenho para sinais e marcações de segurança.
- [2] IDONIC. Contagem de Clientes com CCTV. [ONLINE] – Março/2017.
<http://www.idonic.com/contagem-de-clientes-com-cctv/>
- [3] SNITCH. Contador de Objetos/Pessoas. [ONLINE] – Março/2017.
<http://www.snitch.com.br/cameras-de-seguranca/contador/>
- [4] Bosch Security Systems. Airport. [ONLINE] – Março/2017.
<http://www.boschsecurity.com/corporate/solutions/solutions-detail/airport.html>
- [5] BNP. Bangna Phaisan. Video Surveillance and Recording System.
[ONLINE] - Março/2017.
<http://www.phaisan.com/video-surveillance>
- [6] 2020 VISION TECHNOLOGY. Facial Recognition CCTV. [ONLINE] – Março/2017.
<http://www.2020cctv.com/security-solutions/video-analytics/>
- [7] Vision-Box. Security. [ONLINE] – Março/2017.
<http://www.vision-box.com/solutions/security/>
- [8] Seymen, O., Bolat, O. (2010). The Role of National Culture in Establishing an Efficient Safety Culture in Organizations: An Evaluation in Respect of Hofstede's Cultural Dimensions.
- [9] Carvalho, R. (2011). *Adaptação do Safety Attitudes Questionnaire para o Brasil – Questionário de Atitudes de Segurança*. Ribeirão Preto: Universidade de São Paulo Escola de Enfermagem de Ribeirão Preto.
- [10] Guldenmund, F. (2010). (Mis)understanding Safety Culture and Its relationship to Safety Management. *Risk Analysis*, 10: 1466-1480.
- [11] Health and Safety Executive (HSE). Factors Briefing – Safety Culture. No.7.
- [12] Mayze, B., Bradley, L. (2008) Safety culture: a multilevel assessment tool for the construction industry. Third International Conference of the Cooperative Research Centre (CRC) for Construction Innovation – Clients Driving Innovation: Benefiting from Innovation, Gold Coast, Queensland.

- [13] Neal, A., Griffin, M. (2006). A Study of the Lagged Relationships Among Safety Climate, Safety Motivation, Safety Behavior, and Accidents at the Individual and Group Levels. *Journal of Applied Psychology*, 4: 946–953.
- [14] Machado, M. (2008). *Análise da contribuição do Programa de Segurança Comportamental na prevenção de acidentes de trabalho: um estudo de caso numa empresa química*. Centro Universitário CENAC. São Paulo.
- [15] Health and Safety Executive (HSE). (2005, Outubro). Factores humanos na gestão de perigos de acidentes graves. p.60.
- [16] CITE. Comissão para a Igualdade no Trabalho e no Emprego. Legislação. Legislação Nacional. Código do Trabalho. [ONLINE] – Abril/2017.
http://www.cite.gov.pt/pt/legis/CodTrab_L1_002.html
- [17] IEC/EN 61496 - Segurança de máquinas – Dispositivos de proteção com ação sem contato, “Safety of machinery – Electro-sensitive protective equipment - Part 1: General requirements and tests”, Edition 3.0, International Standard, 2012-04-05.
- [18] Directiva Máquinas 2006/42/CE – Norma Europeia para construção de máquinas. “Diretiva 2006/42/CE Do Parlamento Europeu E Do Conselho de 17 de Maio de 2006 relativa às máquinas e que altera a Directiva 95/16/CE (reformulação)”, Parlamento Europeu, 9.6.2006.
- [19] IEC 62061 - Segurança funcional dos sistemas de controlo eletrónico, eletrónico e programável, “Safety of machinery – Functional safety of safety-related electrical, electronic and programmable electronic control systems”, International Electrotechnical Commission, Edition 1.0 2005-01.
- [20] ISO 13849 - Segurança das máquinas - Peças relacionadas à segurança dos sistemas de controlo, “Safety of machinery - Safety-related parts of control systems - Part 1: General principles for design”, Edition 3, ISO/TC 199 Safety of machinery, 2015-12.
- [21] Changes to IEC 61496 and EN 61496-1, and their consequences for using Electro-Sensitive Protective Devices, Hans-Jörg Stubenrauch, Andreas Sixt.
- [22] IEC 61508 - Segurança Funcional de Sistemas elétricos, eletrónicos, programáveis de Segurança eletrónica (E/E/PE, ou E/E/PES), “Functional safety of electrical/electronic/programmable electronic safety-related systems – Part 1: General requirements”, International Electrotechnical Commission, 1998-12.
- [23] EN 954 - Segurança de máquinas. Peças relacionadas à segurança dos sistemas de controlo, “Safety of machinery. Safety related parts of control systems. General principles for design”, 1997-6-15

- [24] *Electro-sensitive protective devices (ESPE) for safe machines Opto-electronic protective devices*, 2013 - 03, Otto Goernemann, Hans-Joerg Stubenrauch.
- [25] EN ISO 13855 - Segurança das máquinas - Posicionamento de salvaguardas em relação às velocidades de aproximação de partes do corpo humano.
- [26] Pilz. Safe Camera System SafetyEYE. [ONLINE] – Abril/2017.
http://brochures.pilz.nl/bro_pdf/SafetyEYE_2014.pdf.
- [27] Pilz. Safe Camera System SafetyEYE. [ONLINE] – Abril/2017.
<https://www.pilz.com/pt-PT/eshop/00106002207042/SafetyEYE-Safe-camera-system>.
- [28] Direct Industry. Pilz. [ONLINE] – Abril/2017.
<http://www.directindustry.com/pt/prod/pilz/product-7550-1715481.html>
- [29] Newton. QuadCam 3D Zone Monitoring System. [ONLINE] – Abril/2017.
http://www.newtonlabs.com/machviz_quadcam.html
- [30] “Contribuição para o mapeamento tridimensional usando imagem”, Rui Alberto Rodrigues Gomes, 2013.
- [31] Kyto, M, Nuutinen, M, and Oittinen, P, “*Method for measuring stereo camera depth accuracy based on stereoscopic vision*”, SPIE/IS&T *Electronic Imaging Three-Dimensional Imaging, Interaction, and Measurement*. Vol. 7864, 19, 2011.
- [32] Depth Biomechanics. Centre for Sports Engineering Research, Sheffield Hallam University. [ONLINE] – Abril/2017.
<http://www.depthbiomechanics.co.uk/wp-content/uploads/2012/06/stereo-vision-cams.jpg>
- [33] Elphel Inc. Electronics photo-electronics. [ONLINE] – Abril/2017.
http://www3.elphel.com/sites/default/files/stereo_big.jpg
- [34] MTI Instruments. Laser Triangulation. [ONLINE] – Abril/2017.
<http://www.mtiinstruments.com/products/lasertriangulation.aspx>.
- [35] *Radio Detection and Ranging*. [ONLINE] – Abril/2017.
<https://en.wikipedia.org/wiki/Radar>
- [36] *Light Detection and Ranging*. [ONLINE] – Abril/2017.
<https://en.wikipedia.org/wiki/Lidar>

- [37] Gokturk *et al.*, 2004, Gkturk, S. B. ant Tomassi, C. (2004). 3D Head Tracking Based on Recognition and Interpolation Using a Time-Of-Flight Depth Sensor. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Washington D.C., USA*
- [38] Sergi Foix, Guillem Alenya, and Carme Torras. Lock-in time-of-flight (tof) cameras: A survey, *IEEE Sensors Journal*, 11(9); 1917-1926 2011.
- [39] Jan W. Weingarten, Gabriel Grener, and Roland Siegwart. A state-of-the-art 3d sensor for robot navigation. In *In IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*. Pages 2155-2160, 2004.
- [40] Sergi Foix, Guillem Alenya, and Carme Torras. Exploitation of time-of-flight (tof) cameras. Technical report, Institut de Robotica i Informàtica Industrial (IRI), 2007.
- [41] Microsoft. Developer. Kinect for Windows Product Blog. [ONLINE] – Abril/2017.
<https://blogs.msdn.microsoft.com/kinectforwindows/2014/06/05/pre-order-your-kinect-for-windows-v2-sensor-starting-today/>
- [42] Microsoft. Microsoft kinect. [ONLINE] – Abril/2017.
<http://www.xbox.com/en-US/xbox-one/accessories/kinect-for-xbox-one>
- [43] D. Pagliari and L. Pinto, "Calibration of Kinect for Xbox One and Comparison between the Two Generations of Microsoft Sensors," *Sensors*, vol. 15, no. 11, pp. 27569-27589, 2015.
- [44] E. Lachat, H. Macher, T. Landes, and P. Grussenmeyer, "Assessment and Calibration of a RGB-D Camera (Kinect v2 Sensor) Towards a Potential Use for Close-Range 3D Modeling," *Remote Sensing*, vol. 7, no. 10, pp. 13070--13097, 2015.
- [45] J. Sell and P. O'Connor, "The Xbox One system on a chip and Kinect sensor," *IEEE Micro*, no. 2, pp. 44-53, 2014.
- [46] Zugara. [ONLINE] – Abril/2017.
<http://zugara.com/wp-content/uploads/Kinect-1-vs-Kinect-2-Tech-Comparison.png>
- [47] University of Tartu Faculty of Science and Technology Institute of Technology Computer Engineering Curriculum, "3D reconstruction using Kinect v2 camera", Gholamreza Anbarjafari, Morteza Daneshmand, MSc, 2016.
- [48] C. S. Bamji, P. O'Connor, T. Elkhatib, S. Mehta, B. Thompson, L. A. Prather, D. Snow, O. C. Akkaya, A. Daniel, A. D. Payne, et al., "A 0.13 um CMOS system-on-chip for a 512 x 424 time-of-flight image sensor with multi-frequency photo-demodulation up to 130 MHz and 2 GS/s ADC," *Solid-State Circuits, IEEE Journal of*, vol. 50, no. 1, pp. 303-319, 2015.

- [49] F. Jaremo Lawin, “*Depth data processing and 3D reconstruction using the Kinect v2*,” Master's thesis, Linköping University, 2015.
- [50] H. Sarbolandi, D. Leoch, and A. Kolb, “*Kinect range sensing: Structured-light versus Time-of-Flight Kinect*,” *Computer Vision and Image Understanding*, vol. 139, pp. 1-20, 2015.
- [51] Microsoft, Kinect SDK. [ONLINE] – Maio/2017.
<https://developer.microsoft.com/en-us/windows/kinect>
- [52] P. Sturm, “Pinhole camera model,” in *Computer Vision*, pp. 610-613, Springer, 2014.
- [53] Microsoft, Camera intrinsics structure. [ONLINE] – Maio/2017.
https://github.com/Kinect/Docs/blob/master/Kinect4Windows2.0/k4w2/Reference/Kinect_for_Windows_v2/Kinect/CameraIntrinsics_Structure.md
- [54] Z. Zhang, “A exible new technique for camera calibration,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 22, no. 11, pp. 1330-1334, 2000
- [55] C. Zhang and Z. Zhang, “Calibration between depth and color sensors for commodity depth cameras,” in *Computer Vision and Machine Learning with RGB-D Sensors*, pp. 47-64, Springer, 2014.
- [56] Mittet, M.-A., Grussenmeyer, P., Landes, T., Yang, Y. and Bernard, N., 2013. Mobile outdoor relative localization using calibrated RGB-D cameras. In: *8th International Symposium on Mobile Mapping Technology*, pp. 2-7.
- [57] E. Lachat, H. Macher, M.-A. Mittet, T. Landes, P. Grussenmeyer, “First Experiences with Kinect V2 Sensor for Close Range 3D Modelling”, Photogrammetry and Geomatics Group, ICube Laboratory UMR 7357, INSA Strasbourg, France, 2015.
- [58] Wikipédia. Microsoft Visual Studio. [ONLINE] – Maio/2017.
https://pt.wikipedia.org/wiki/Microsoft_Visual_Studio
- [59] Wikipédia. Matlab. [ONLINE] – Maio/2017.
<https://pt.wikipedia.org/wiki/MATLAB>
- [60] MathWorks. Documentation. Delaunay Triangulation. [ONLINE] – Maio/2017.
<https://www.mathworks.com/help/matlab/math/delaunay-triangulation.html>
- [61] Satyan L. Devadoss and Joseph O'Rourke *Discrete and Computational Geometry*, Princeton University Press, 2011.
- [62] Mark de Berg, Otfried Cheong, Marc van Kreveld, Mark Overmars, *Computational Geometry: Algorithms and Applications*, Springer, 3rd edition, 2008.

- [63] Franco P. Preparata, Michael I. Shamos, *Computational Geometry: An Introduction*, Springer, 1985.
- [64] O'Rourke, J., *Computational Geometry in C*, Cambridge University Press, 1994.
- [65] Snoeyink, J., *Pointlocation*, in *Handbook of Discrete and Computational Geometry*, Third Edition edited by Jacob E. Goodman, Joseph O'Rourke, and Csaba D. Tóth CRC Press LLC, 2017.
- [66] Arduino. *Arduino Uno & Genuino Uno*. [ONLINE] – Maio/2017.
<https://www.arduino.cc/en/main/arduinoBoardUno>
- [67] MathWorks. *Hardware Support*. *Arduino Support from Matlab*. [ONLINE] – Maio/2017.
<https://www.mathworks.com/hardware-support/arduino-matlab.html>
- [68] “*Evaluating and Improving the Depth Accuracy of Kinect for Windows v2*”, Longyu Zhang, Abdulmotaleb El Saddik, Article in *IEEE Sensors Journal* · August 2015.

ANEXOS

“Anexo I

Requisitos mínimos dos sistemas de videovigilância

[a que se refere a alínea *a*) do n.º 1 do artigo 7.º]

1. Requisitos técnicos mínimos:

1.1. Os requisitos técnicos mínimos das câmaras de videovigilância são definidos de acordo com os fins a que se destina a videovigilância nos termos do disposto do n.º 1 do artigo 31.º da Lei n.º 34/2013, de 16 de maio, com as seguintes especialidades:

a) Para proteção de edifícios e respetivos acessos, devem as câmaras:

- i) Ser, preferencialmente, policromáticas;
- ii) Ter capacidade de iluminação, resolução, ampliação e abertura de foco que garanta a deteção e reconhecimento inequívoco de vultos tipo humano;

b) Para proteção de instalações em que sejam estabelecidos requisitos especiais de proteção nos termos previstos na presente portaria, devem as câmaras:

- i) Ser policromáticas;
- ii) Permitir a gravação de som quando autorizada pela CNPD;
- iii) Ter capacidade de iluminação, resolução, ampliação e abertura de foco que garanta o reconhecimento e a identificação de indivíduos, podendo, para o efeito, ser tomado como referência o disposto no Regulamento (CE) n.º 2252/2004, do Conselho, de 13 de dezembro (estabelece normas para os dispositivos de segurança e dados biométricos dos passaportes e documentos de viagem emitidos pelos Estados membros);

1.2. Para além dos requisitos específicos enunciados em 1.1., todas as câmaras de videovigilância devem ainda garantir:

- a) A proteção contra vandalismo e índice de proteção compatível com o IP66;
- b) O uso dos sistemas normalizados de compressão de acordo com a norma H264 ou equivalente;
- c) A definição lógica ou a utilização física de máscaras nos locais em que é legalmente proibida a captação de imagens.

1.3. Os requisitos técnicos mínimos de comunicação são:

- a) A transmissão de imagens, bem como de som quando legalmente autorizada, bem como o controlo e gestão das câmaras;
- b) A utilização de sistemas sem fios para transmissão de dados que tem de ser feita de ponto a ponto, a título de exclusividade da câmara;
- c) Todas as transmissões são encriptadas, tendo a chave de encriptação de ser alterada a cada seis meses.

2. Visualização e monitorização

O sistema de controlo deve ser operado em ambiente seguro e deve garantir:

- a) A visualização, o controlo e a gestão das câmaras de videovigilância em tempo real;
- b) A autenticação dos operadores.

3. Registos de segurança e auditorias

3.1. A gravação local ou remota das imagens pelas câmaras de videovigilância é feita:

- a) Em formato digital;
- b) De forma encriptada;
- c) Em tempo real, tendo os servidores de estar sincronizados com a hora legal portuguesa, por forma a garantir a fidedignidade da data e hora que devem constar de cada imagem captada;
- d) De forma a que seja auditável.

3.2. Todas as intervenções realizadas ao nível dos sistemas locais são registadas:

- a) Em formato digital;
- b) De forma encriptada;
- c) Em tempo real;
- d) De forma a que sejam auditáveis.

3.3. A operação do sistema local requer obrigatoriamente que o sistema de registo de eventos esteja ativo, a fim de garantir as operações de auditoria.

4. A adaptação dos sistemas já em funcionamento deve ter lugar no prazo de 2 anos.”

“Anexo VIII

Sinalização de meios de vigilância eletrónica

(a que se refere o n.º 1 do artigo 115.º)

Sinal em forma de triângulo equilátero, em fundo de cor amarela com orla interior em cor preta, ao centro, símbolo representando o pictograma de uma câmara de videovigilância em cor preta.

Símbolo gráfico”

