

Πολυτεχνείο Κρήτης  
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών



Τρισδιάστατη Ανακατασκευή Αντικειμένων σε  
Πραγματικό Χρόνο μέσω Ενσωματωμένου Συστήματος

ΔΙΠΛΩΜΑΤΙΚΉ ΕΡΓΑΣΙΑ

Βενιαμίν Μαλεφίουδάκης

Επιτροπή:

Μιχάλης Ζερβάκης, Καθηγητής (Επιβλέπων)  
Απόστολος Δόλλας, Καθηγητής  
Αριστομένης Αντωνιάδης, Καθηγητής

Χανιά, Ιούνιος 2022



## Περίληψη

Η τρισδιάστατη ανακατασκευή αντικειμένων εξελίσσεται ολοένα και περισσότερο στο πεδίο της μηχανικής όρασης. Μέσω κατάλληλων αλγορίθμων, υπάρχει πλέον η δυνατότητα να παραχθούν συνθέτες τρισδιάστατες σκηνές σε πραγματικό χρόνο. Σε αυτή την κατεύθυνση, η σύνθεση ρεαλιστικών τρισδιάστατων μοντέλων παραμένει απαιτητική και χρονοβόρα, αυξάνοντας σημαντικά το υπολογιστικό κόστος. Στη παρούσα διπλωματική εργασία, βασικό αντικείμενο αποτέλεσε η ανάπτυξη και υλοποίηση μιας ολοκληρωμένης εφαρμογής για τη τρισδιάστατη ανακατασκευή εικόνων (χάρτης ανισοτιμίας – disparity map) σε ενσωματωμένο σύστημα, με τη χρήση οπτικών αισθητήρων πολύ χαμηλού κόστους. Η εφαρμογή σχεδιάστηκε και υλοποιήθηκε με σκοπό να διαθέτει γραφικό περιβάλλον που να επιτρέπει στο χρήστη να εκτελέσει εύκολα όλες τις διαδικασίες που απαιτούνται για τη τρισδιάστατη ανακατασκευή της σκηνής. Μπορεί επίσης να προσφέρει τη δυνατότητα επιλογής ενός εκ των δύο στερεοσκοπικών αλγορίθμου που ενσωματώθηκαν για τη δημιουργία του χάρτη βάθους της σκηνής, καθώς και τη ρύθμιση των παραμέτρων τους σε πραγματικό χρόνο. Η εφαρμογή έχει αναπτυχθεί με βέλτιστο τρόπο ως προς την αποτελεσματική αξιοποίηση της υπολογιστή ισχύς του συστήματος. Ο χρήστης μπορεί να επιλέξει είτε την υπολογιστική μονάδα επεξεργασίας (CPU) ή τη μονάδα επεξεργασίας γραφικών (GPU) για τον υπολογισμό των τρισδιάστατων γραφικών σε πραγματικό χρόνο. Οι υπολογισμοί και συγκρίσεις που πραγματοποιήθηκαν σε αυτή τη διπλωματική εστιάζονται στη χρήση συμβατού σταθερού υπολογιστή (Desktop PC) και στη χρήση του ενσωματωμένου συστήματος Nvidia Jetson NX Xavier (Jetson). Οι συγκρίσεις μεταφράζονται σε όρους χρονικών υπολογισμών με είσοδο εικόνες των 640 ή 1080 Pixel, μεταξύ των δύο διαφορετικών αλγορίθμων, μεταξύ των CPU και GPU και μεταξύ των δύο υπολογιστικών συστημάτων (Desktop PC vs Jetson). Από αυτές τις συγκρίσεις, παρουσιάζεται ότι η χρήση Jetson για τον υπολογισμό disparity map σε πραγματικό χρόνο είναι πάρα πολύ κοντά στους χρόνους του Desktop PC (διαφορά 8 ms για υπολογισμούς σε GPU και χρήση 640 Pixel και 100 ms για υπολογισμούς σε CPU και χρήση 1080 Pixel). Μέσα από αυτά τα αποτελέσματα, γίνεται αντιληπτό ότι η χρήση ενσωματωμένων συστημάτων με χαμηλού κόστους οπτικούς αισθητήρες είναι πλέον σε θέση να μπορούν να υποστηρίξουν φορητά και επανδρωμένα ή (μη) συστήματα με ένα ευρύ φάσμα εφαρμογής όπως είναι η εναέρια ανακατασκευή τρισδιάστατου χώρου.

# Abstract

The three-dimensional reconstruction of objects is evolving more and more in the field of machine vision. Through appropriate algorithms, it is now possible to produce synthetic 3D scenes in real time. In this direction, the synthesis of realistic 3D models remains demanding and time-consuming, significantly increasing the computational cost. In this thesis, the main purpose is the development and implementation of a integrated application for 3D image reconstruction (disparity map) in an embedded system, using optical sensors at very low cost. The application is designed and implemented to have a graphical interface that allows the user to easily perform all the procedures required for the 3D reconstruction of the scene. In addition, it offers the possibility of selecting one of the two stereoscopic algorithms incorporated to create the depth map of the scene and to adjust their parameters in real time. The application has been optimally developed in terms of the efficient use of computer system power. The user can select either the computing processing unit (CPU) or the graphics processing unit (GPU) for the calculation of the 3D graphics in real time. The calculations and comparisons performed in this thesis are focused on the use of a compatible desktop PC and the use of the embedded system Nvidia Jetson NX Xavier (Jetson). Comparisons with input images of 640 or 1080 pixels have been made, between the two different algorithms, between the CPU and GPU and between the two computing systems (Desktop PC vs. Jetson). From these comparisons, it is shown that using Jetson for real-time disparity map computation is very close to the Desktop PC times (8 ms difference for GPU calculations and using 640 pixels and 100 ms for CPU calculations using 1080 pixels). From these results, it can be seen that the use of embedded systems with low optical sensors are now able to support portable and manned or unmanned systems with a wide range of applications such as airborne 3D space reconstruction.

## *Ευχαριστίες*

*Σε αυτό το σημείο θα ήθελα να ευχαριστήσω τον επιβλέποντα Καθηγητή Μιχάλη Ζερβάκη για την καθοδήγηση και την πολύτιμη βοήθεια, που μου προσέφερε κατά τη διάρκεια της εκπόνησης της διπλωματικής εργασίας. Επιπλέον θα ήθελα να ευχαριστήσω τα υπόλοιπα μέλη της επιτροπής τον Καθηγητή Απόστολο Δόλλα και τον Καθηγητή Αριστομένη Αντωνιάδη.*

*Επίσης θα ήθελα να ευχαριστήσω τον διδάκτορα Μάριο Αντωνακάκη, για την αρωγή του σε οποιοδήποτε σημείο χρειάστηκα βοήθεια, καθώς και για τη στήριξη που μου έδειξε σε όλη την πορεία της διπλωματικής εργασίας.*

*Τέλος, θα ήθελα να ευχαριστήσω την οικογένεια μου που με πιστεύει και με στηρίζει με κάθε δυνατό τρόπο.*



# Περιεχόμενα

<b>1</b>	<b>Εισαγωγή</b>	<b>10</b>
1.1	Κίνητρο . . . . .	10
1.2	Περιγραφή της εργασίας . . . . .	11
1.3	Σύντομη βιβλιογραφική ανασκόπηση . . . . .	12
<b>2</b>	<b>Θεωρητικό υπόβαθρο</b>	<b>13</b>
2.1	Μηχανική όραση (Computer vision) . . . . .	13
2.2	Στερεοσκοπική όραση (Stereo vision) . . . . .	14
2.3	Στερεοσκοπικό σύστημα (Stereo vision system) . . . . .	14
2.4	Βασικά βήματα για την μοντελοποίηση της 3D σκηνής μέσω του στερεοσκοπικού συστήματος . . . . .	15
2.5	Βαθμονόμηση κάμερας (Camera calibration) . . . . .	17
2.5.1	Συστήματα συντεταγμένων . . . . .	17
2.5.2	Μοντέλο κάμερας (Camera model) . . . . .	18
2.5.3	Προοπτική προβολή (Perspective transform) . . . . .	20
2.5.4	Ενδογενείς παράμετροι (Intrinsic parameters) . . . . .	22
2.5.5	Εξωγενείς παράμετροι (Extrinsic parameters) . . . . .	23
2.5.6	Παραμορφώσεις (Distortions) . . . . .	24
2.5.7	Ακτινική παραμόρφωση(Radial distortion) . . . . .	25
2.5.8	Εφαπτομενική παραμόρφωση (Tangential distortion) . . . . .	28
2.5.9	Μέθοδοι βαθμονόμησης(Calibration methods) . . . . .	29
2.5.10	Ακρίβεια (Accuracy) . . . . .	32
2.6	Ανόρθωση εικόνας (Image rectification) . . . . .	34
2.6.1	Επιπολική γεωμετρία (Epipolar geometry) . . . . .	35
2.6.2	Στερεοσκοπική ανόρθωση (Stereo rectification) . . . . .	38
2.6.3	Αλγόριθμος Bouguet (Bouguet's algorithm) . . . . .	40
2.7	Στερεοσκοπική αντιστοίχιση (Stereo matching) . . . . .	42
2.7.1	Αλγόριθμοι στερεοσκοπικής αντιστοίχισης (Stereo matching algorithm) . . . . .	42
2.7.2	Προβλήματα στερεοσκοπικής αντιστοίχιας . . . . .	44
2.7.3	Κατηγορίες αλγόριθμων στερεοσκοπικής συσχέτισης . . . . .	47
2.7.4	Stereo Block Matching . . . . .	47
2.7.5	Semi Global Block Matching . . . . .	49
2.8	Ανακατασκευή σκηνής - Τριγωνοποίηση . . . . .	51
<b>3</b>	<b>Υλοποίηση εφαρμογής</b>	<b>54</b>
3.1	Βιβλιοθήκη OpenCV . . . . .	54
3.2	Compute Unified Device Architecture (CUDA) . . . . .	55

3.3	Βιβλιοθήκη Qt . . . . .	56
3.4	Qt Creator . . . . .	57
3.5	Υλοποίηση εφαρμογής βαθμονόμησης . . . . .	59
3.6	Μέθοδος ανόρθωσης (Rectification method) . . . . .	64
3.6.1	InitUndistortRectifyMap . . . . .	64
3.6.2	Remap . . . . .	65
3.6.3	Υλοποίηση στερεοσκοπικής ανόρθωσης . . . . .	65
3.7	Υλοποίηση στερεοσκοπικής συσχέτισης . . . . .	67
<b>4</b>	<b>Περιγραφή εξοπλισμού</b>	<b>71</b>
4.1	Περιγραφή ενσωματωμένου συστήματος . . . . .	71
4.2	Προδιαγραφές καμερών . . . . .	72
4.3	Υλοποίηση στερεοσκοπικού συστήματος . . . . .	73
<b>5</b>	<b>Αποτέλεσμα</b>	<b>74</b>
5.1	Αποτελέσματα βαθμονόμησης στερεοσκοπικού συστήματος . . . . .	74
5.2	Αποτελέσματα ανόρθωσης εικόνων . . . . .	78
5.3	Αποτελέσματα αλγορίθμων στερεοσκοπικής συσχέτισης . . . . .	79
<b>6</b>	<b>Συμπεράσματα και προτάσεις για μελλοντική έρευνα</b>	<b>88</b>
	<b>Βιβλιογραφία</b>	<b>93</b>

## Κατάλογος Σχημάτων

1	Mars Rover [20] . . . . .	14
2	Διατάξεις στερεοσκοπικού συστήματος [18] . . . . .	15
3	Διαδικασία μοντελοποίησης της 3D σκηνής, μέσω του στερεοσκοπικού συστήματος [8] . . . . .	16
4	Συστήματα Αναφοράς. . . . .	18
5	Μοντέλο κάμερας «μικρής οπής» . . . . .	19
6	Προβολή ενός σημείου . . . . .	19
7	Σύστημα συντεταγμένων με το επίπεδο εικόνας να είναι μπροστά από την εστία της κάμερας . . . . .	21
8	Από τις συντεταγμένες του επιπέδου εστίασης στις συντεταγμένες της εικόνας . . . . .	23
9	Παράδειγμα ακτινικής παραμόρφωσης [15] . . . . .	25
10	Παράδειγμα θετικής ακτινικής παραμόρφωσης . . . . .	26
11	Παράδειγμα αρνητικής ακτινικής παραμόρφωσης . . . . .	26
12	Η αίτια της ακτινικής παραμόρφωσης[8] . . . . .	27
13	Εφαπτομενική παραμόρφωση . . . . .	28
14	Τρισδιάστατο αντικείμενο παρατήρησης [16] . . . . .	30
15	Δισδιάστατο αντικείμενο παρατήρησης . . . . .	30
16	Μονοδιάστατο αντικείμενο παρατήρησης[16] . . . . .	31
17	Στερεοσκοπική ανόρθωση εικόνων. [8] . . . . .	34
18	Επιπολική γεωμετρία[8] . . . . .	35
19	Επιπολικές γραμμές [13] . . . . .	37
20	Στερεοσκοπική ανόρθωση [8] . . . . .	39
21	Στερεοσκοπική ανόρθωση αποτελέσματα [8] . . . . .	40
22	Πρόβλημα στερεοσκοπικής αντιστοίχισης: Φωτομετρική παραμόρφωση [16] . . . . .	44
23	Πρόβλημα στερεοσκοπικής αντιστοίχισης: Προοπτική παραμόρφωση [16] . . . . .	45
24	Πρόβλημα στερεοσκοπικής αντιστοίχισης: Ψευδείς αντιστοιχίες [16] . . . . .	45
25	Πρόβλημα στερεοσκοπικής αντιστοίχισης: Αποκλεισμένες περιοχές [16] . . . . .	46
26	Εφαρμογή τριγωνοποίησης για τον υπολογισμό του βάθους [13] . . . . .	51
27	Disparity map [13] . . . . .	52
28	Ροή δεδομένων CPU to GPU . . . . .	55
29	Qt: Singals and Slots . . . . .	56
30	Qt Creator: Menu GUI . . . . .	57
31	Qt Creator: Disparity Map GUI . . . . .	58
32	Εντολή στο terminal για την τοποθεσία των καμερών στο υπολογιστή . . . . .	60
33	Εφαρμογή για την βαθμονόμηση της στερεοσκοπικής κάμερας. . . . .	60
34	Εφαρμογή για φωτογράφιση του αντικειμένου βαθμονόμησης. . . . .	61

35	Εφαρμογή για φωτογράφιση του αντικειμένου βαθμονόμησης. . . . .	61
36	Ενδεικτικές φωτογραφίες της σκακιέρας. [8] . . . . .	62
37	Ενδεικτικές φωτογραφίες της σκακιέρας . . . . .	63
38	Εφαρμογή για φωτογράφιση του αντικειμένου βαθμονόμησης. . . . .	63
39	Μενού Εφαρμογής, στερεοσκοπική ανόρθωση . . . . .	66
40	Αποτελέσματα της στερεοσκοπικής ανόρθωσης μέσω της εφαρμογής . .	66
41	Εφαρμογή: Disparity map . . . . .	68
42	Ενσωματωμένο σύστημα: Jetson Xavier NX. [11] . . . . .	71
43	Εικόνα κάμερας . . . . .	72
44	Στερεοσκοπική διάταξη καμερών . . . . .	73
45	Αντικείμενο βαθμονόμησης τύπου σκακιέρας . . . . .	74
46	Ενδεικτικές λήψεις του αντικειμένου βαθμονόμησης . . . . .	75
47	Αρχικές εικόνες από το στερεοσκοπικό σύστημα . . . . .	78
48	Εικόνες από το στερεοσκοπικό σύστημα μετά την εφαρμογή της στερεο- σκοπικής ανόρθωσης . . . . .	78
49	Διαμόρφωση παραμέτρων του αλγορίθμου Stereo Block Matching ( <i>CPU</i> ) . . . . .	80
50	Disparity map αλγορίθμου Stereo Block Matching ( <i>CPU</i> ) . . . . .	80
51	Διαμόρφωση παραμέτρων του αλγορίθμου Stereo Block Matching ( <i>CUDA</i> ) . . . . .	81
52	Disparity map αλγορίθμου Stereo Block Matching ( <i>CUDA</i> ) . . . . .	82
53	Point Cloud αλγορίθμου Stereo Block Matching ( <i>CUDA</i> ) μέσω MeshLab	83
54	Διαμόρφωση παραμέτρων του αλγορίθμου Semi Global Stereo Matching	84
55	Disparity map αλγορίθμου Semi Global Stereo Matching . . . . .	85
56	Point Cloud αλγορίθμου Semi Global Stereo Matching μέσω MeshLab	86

# 1 Εισαγωγή

## 1.1 Κίνητρο

Η τρισδιάστατη απεικόνιση άρχισε να αποκτά όλο ένα και περισσότερο έδαφος στις εξελίξεις της τεχνολογίας. Υπάρχουν περιπτώσεις όπου είναι απαραίτητη η γνώση της τρισδιάστατης δομής ενός αντικείμενου ή μιας σκηνής. Πλέον είναι εφικτό, μέσω κατάλληλων αλγορίθμων, να παραχθούν συνθέτες τρισδιάστατες σκηνές σε πραγματικό χρόνο. Το πρόβλημα είναι ότι παρόλο που η υπολογιστή ισχύς έχει αυξηθεί σημαντικά, η σύνθεση των ρεαλιστικών τρισδιάστατων μοντέλων παραμένει απαιτητική και χρονοβόρα, αυξάνοντας σημαντικά το κόστος. Η επίλυση του παραπάνω προβλήματος έχει απασχολήσει την επιστημονική κοινότητα. Αρκεί μια σύντομη βιβλιογραφική αναζήτηση για να διαπιστώσει κανείς τις διαφορετικές προσεγγίσεις που έχουν αναπτύξει κατά καιρούς οι ερευνητές.

Η εργασία του Chand αναφέρεται στη διαδικασία της τρισδιάστατη ανακατασκευής της σκηνής με τη χρήση ενός στερεοσκοπικού συστήματος. Πιο συγκεκριμένα, περιγράφεται πως μπορεί να πραγματοποιηθεί η βαθμονόμηση ενός στερεοσκοπικού συστήματος, η δημιουργία του χάρτη βάθους της σκηνής και η απεικόνιση της τρισδιάστατης ανακατασκευής, μέσω βιβλιοθηκών ανοιχτού κωδικά. Το βασικό μειονέκτημα είναι ότι, ενώ περιγράφεται αναλυτικά η διαδικασία, είναι σχετικά περίπλοκη και απαιτεί καλές γνώσεις προγραμματισμού. [4]

Οι Ildar και Pomazon προτείνουν ένα ολοκληρωμένο σύστημα χαμηλού κόστους για τον υπολογισμό του χάρτη βάθους της σκηνής. Επίσης, έχουν υλοποιήσει, σε γλώσσα Python, μια ολοκληρωμένη εφαρμογή, για την εύκολη ρύθμιση των παραμέτρων του αλγορίθμου, που εκτελείται σε πραγματικό χρόνο. Περιορισμό της εφαρμογής αυτής αποτελεί η έλλειψη δυνατότητας επιλογής του αλγορίθμου από το χρήστη, μεταξύ αυτών που έχουν αναπτυχθεί για τον σκοπό αυτό. [6]

Οι Κανελλάκης και Κυρίτσης στην εργασία τους ανέπτυξαν μια βιομηχανική διάταξη, η οποία αποτελείται από ένα ρομποτικό βραχίονα και ένα στερεοσκοπικό σύστημα. Αναφέρεται χαρακτηριστικά ότι χρειάστηκαν αρκετές δοκιμές για να επιλέγει ο κατάλληλος αλγόριθμος για τον υπολογισμό του χάρτη βάθους, καθώς και να ρυθμιστούν οι παράμετροι του. [7]

Ο στόχος της παρούσας διπλωματικής εργασίας είναι ο πειραματισμός και η ανάπτυξη εφαρμογής σε ενσωματωμένο σύστημα που διαθέτει μονάδες παράλληλης επεξεργασίας γραφικών – GPU (Nvidia Jetson NX Xavier) για την τρισδιάστατη ανακατασκευή αντικείμενων σε πραγματικό χρόνο με τη χρήση χαμηλού κόστους καμερών. Η πειραματική διαδικασία έχει ως στόχο να αναδείξει κατά πόσο μπορούν να αξιοποιηθούν στο πεδίο της στερεοσκοπικής όρασης ενσωματωμένα συστήματα τελευταίας τεχνολογίας με GPU. Επιπρόσθετα, η ανάπτυξη ενός γραφικού περιβάλλοντος αλληλεπίδρασης με τον χρήστη, για θέματα στερεοσκοπικής όρασης, στοχεύει στην κάλυψη της έλλειψης ανοι-

κτού κώδικά. Η αναπτυσσόμενη εφαρμογή συνδυάζει όλες τις απαραίτητες λειτουργίες, ώστε να γίνει η τρισδιάστατη ανακατασκευή. Επίσης, μέσω του γραφικού περιβάλλοντος που διαθέτει, επιτρέπει στον χρήστη να εκτελεί με ευκολία βασικές διαδικασίες, όπως η βαθμονόμηση των καμερών. Επιπλέον, παρέχει τη δυνατότητα επιλογής του αλγόριθμου στερεοσκοπικής συσχέτισης, καθώς και ρύθμισης των παραμέτρων σε πραγματικό χρόνο. Τέλος, εφόσον έχει δημιουργηθεί ο χάρτης βάθους, γίνεται η παράγωγή 3D point clouds.

Το στερεοσκοπικό σύστημα που προτείνεται αποτελείται από δυο κάμερες χαμηλού κόστους τύπου Web Camera, οι οποίες έχουν τοποθετηθεί με βάση την απλή στερεοσκοπική διάταξη. Η απόσταση μεταξύ τους είναι 10cm. Η διάταξη αυτή προσομοιώνει τη θέση των ανθρώπινων οφθαλμών

## 1.2 Περιγραφή της εργασίας

Η εργασία χωρίζεται σε έξι κεφάλαια. Παρακάτω θα δοθεί μια σύντομη αναφορά του κάθε κεφάλαιου.

Στο **κεφάλαιο 2** αναλύεται το απαραίτητο για τον αναγνώστη θεωρητικό υπόβαθρο. Αρχικά, περιγράφονται σύντομα βασικές έννοιες της μηχανικής όρασης. Στη συνέχεια, δίνεται έμφαση σε σημαντικά εδάφια που είναι αναγκαία για την τρισδιάστατη ανακατασκευή εικόνων, όπως είναι η βαθμονόμηση των καμερών, η ανόρθωση εικόνων, η στερεοσκοπική συσχέτιση, καθώς και η ανακατασκευή των εικόνων.

Στο **κεφάλαιο 3** περιγράφεται η υλοποίηση της εφαρμογής που αναπτύχθηκε στα πλαίσια της διπλωματικής. Πιο συγκεκριμένα, αναλύεται ο σχεδιασμός του γραφικού περιβάλλοντός της, καθώς και οι προϋποθέσεις που απαιτούνται. Επιπλέον, περιγράφεται η κατασκευή του στερεοσκοπικού συστήματος και το ενσωματωμένο σύστημα που χρησιμοποιήθηκε.

Στο **κεφάλαιο 4** αναλύονται τα αποτελέσματα που προέκυψαν από το στερεοσκοπικό σύστημα σε συνδυασμό με την εφαρμογή που αναπτύχθηκε. Επίσης, γίνεται σύγκριση των αποτελεσμάτων μεταξύ του ενσωματωμένου συστήματος με έναν υπολογιστή γραφείου.

Στο **κεφάλαιο 5** παρουσιάζονται τα συμπεράσματα και προτείνονται μελλοντικές βελτιώσεις.

### 1.3 Σύντομη βιβλιογραφική ανασκόπηση

Το Stereo vision είναι μια μέθοδος που χρησιμοποιείται σε διάφορα πεδία της βιομηχανίας και ως εκ τούτου δημοσιεύονται πολλές εργασίες και εφαρμογές πάνω σε αυτό το θέμα. Ένας από τους κυρίους στόχους είναι η μείωση του κόστους χρησιμοποιώντας διάφορες τεχνολογίες και μεθόδους. Έχει πραγματοποιηθεί αρκετή έρευνα για το θέμα αυτό, για παράδειγμα, ο Oleari [12] στην εργασία του παρουσιάζει ένα σύστημα όπου το συνολικό το κόστος είναι 80 ευρώ. Χρησιμοποιεί κάμερες Logitech C270, οι οποίες μπορούν να παρέχουν εικόνες ανάλυσης SVGA ( $800 \times 600$ ) στα 15 *fps*.

Πολλές εργασίες χρησιμοποιούν βιβλιοθήκες ανοιχτού κωδικά για τις υλοποιήσεις τους, μεταξύ αυτών μεγάλη χρήση συναντά η βιβλιοθήκη OpenCV και τα εργαλεία της. Ο Pomaska [17] χρησιμοποιεί το OpenCV σε συνδυασμό με RaspberryPI για να υπολογίσει την απόσταση των αντικείμενων. Παρόμοια δουλεία έχει κάνει και ο Chaohui [9], παρουσιάζοντας επιπλέον ένα γραφικό περιβάλλον για την απλοποίηση της διαδικασίας. Οι Ildar και Pomazon[6] προτείνουν ένα ολοκληρωμένο σύστημα χαμηλού κόστους για τον υπολογισμό του χάρτη βάθους της σκηνής, χρησιμοποιώντας το ενσωματωμένο σύστημα Jetson Nano. Επίσης, έχουν υλοποιήσει, σε γλώσσα Python, μια ολοκληρωμένη εφαρμογή, για την εύκολη ρύθμιση των παραμέτρων του αλγορίθμου, που εκτελείται σε πραγματικό χρόνο.

Κάποιες σχετικές εργασίες προτείνουν τη χρήση νευρωνικών δικτύων για τον υπολογισμό του χάρτη βάθους, χρησιμοποιώντας μονό μια κάμερα. Σε αυτή την περίπτωση αυξάνεται η απαιτούμενη υπολογιστική ισχύς, με αποτέλεσμα να είναι δύσκολο να εκτελεστούν άλλες διεργασίες.[22]. Επίσης, χρησιμοποιούνται διάφορες τεχνολογίες για τη στερεοσκοπική αντιστοιχία, όπως για παράδειγμα, στην εργασία [10] προτείνουν ένα χαμηλού κόστους σύστημα χρησιμοποιώντας την τεχνολογία FPGS.

## 2 Θεωρητικό υπόβαθρο

### 2.1 Μηχανική όραση (Computer vision)

Η μηχανική όραση, υπολογιστική όραση ή τεχνητή όραση είναι ένα επιστημονικό πεδίο της τεχνητής νοημοσύνης, το οποίο επιχειρεί να αναπαράγει αλγοριθμικά την αίσθηση της όρασης, συνήθως σε ηλεκτρονικό υπολογιστή ή ρομπότ. Η μηχανική όραση σχετίζεται με τη θεωρία και την τεχνολογία που εμπλέκονται στη σχεδίαση και κατασκευή συστημάτων που λαμβάνουν και αναλύουν δεδομένα από ψηφιακές εικόνες. [19]

Η τρισδιάστατη απεικόνιση άρχισε να αποκτά όλο ένα και περισσότερο έδαφος στις εξελίξεις της τεχνολογίας. Υπάρχουν περιπτώσεις όπου είναι απαραίτητη η γνώση της τρισδιάστατης δομής ενός αντικείμενου ή μιας σκηνής. Πλέον είναι εφικτό, μέσω κατάλληλων αλγορίθμων, να παραχθούν συνθέτες τρισδιάστατες σκηνές σε πραγματικό χρόνο. Το πρόβλημα είναι ότι παρόλο που η υπολογιστή ισχύς έχει αυξηθεί σημαντικά, η σύνθεση των ρεαλιστικών τρισδιάστατων μοντέλων παραμένει απαιτητική και χρονοβόρα με αποτέλεσμα το υψηλό κόστος.

Τα τελευταία χρόνια έχουν αναπτυχθεί αρκετές τεχνικές για την τρισδιάστατη ανακατασκευή αντικείμενων. Οι τεχνικές αυτές χωρίζονται σε δυο κατηγορίες, τις ενεργές και τις παθητικές. Οι ενεργές μέθοδοι χρησιμοποιούν ειδικές συσκευές, οι οποίες μέσω της ενέργειας προσπαθούν να εκτιμήσουν την απόσταση κάθε σημείου του αντικείμενου. Στις παθητικές τεχνικές γίνεται χρήση ενός μέσου για την καταγραφή του αντικείμενου και στη συνέχεια πραγματοποιείται επεξεργασία των δεδομένων αυτών.

Μέσω των παθητικών μεθόδων εξάγονται από την εικόνα πληροφορίες όπως είναι το βάθος. Γενικά, για την τρισδιάστατη αναπαράσταση ενός αντικείμενου χρειάζονται εικόνες του από διαφορετικές λήψεις. Με αυτόν τον τρόπο, είναι εφικτό να συλλεχθούν χρήσιμα στοιχεία όπως το βάθος, η απόσταση των αντικειμένων που απεικονίζονται, οι ακμές, οι μεταβολές του φωτισμού, η υφή και άλλα. Κάποιες από τις βασικές παθητικές τεχνικές για τον υπολογισμό της απόστασης είναι οι εξής: [15]

- **Στερεοσκοπική όραση** (Stereo vision).
- **Δομή από κίνηση** (Structure from motion).
- **Σχήμα από σκίαση** (Shape from shading).
- **Βάθος από μη εστίαση** (Depth from defocus).
- **Βάθος από εστίαση** (Range from focus).
- **Σχήμα από υφή** (Shape from texture).

Η παρούσα διπλωματική εργασία επικεντρώνεται στη μέθοδο της στερεοσκοπικής όρασης.

## 2.2 Στερεοσκοπική όραση (Stereo vision)

Η στερεοσκοπική όραση είναι μια παραδοσιακή μέθοδος για την απόκτηση τρισδιάστατων πληροφοριών από ένα στερεοσκοπικό ζεύγος εικόνων, κατά την οποία αποτυπώνονται δύο διαφορετικές λήψεις του ίδιου αντικείμενου ή της ίδιας σκηνής.

Με μια μόνο εικόνα είναι πολύ δύσκολο να αποκτηθούν πληροφορίες για την τρισδιάστατη δομή της σκηνής, έκτος αν γίνεται εφαρμογή κατάλληλων αισθητήρων. Ο λόγος είναι ότι κατά τη λήψη της εικόνας χάνεται η αίσθηση του βάθους. Όμως με τη μέθοδο αυτή γίνεται εφικτή η εκτίμηση της τρισδιάστατης δομής και της απόστασης των αντικειμένων. Η τεχνική αυτή χρησιμοποιείται σε πολλές εφαρμογές λόγω των πολλών πλεονεκτημάτων που παρέχει. Μεταξύ άλλων αναφέρονται κάποια από τα πλεονεκτήματα όπως το κόστος, η ασφάλεια, το εύρος λειτουργίας και η αξιοπιστία.

## 2.3 Στερεοσκοπικό σύστημα (Stereo vision system)

Ένα στερεοσκοπικό ζεύγος εικόνων αποτελείται από δυο κάμερες που έχουν τοποθετηθεί κατάλληλα, ώστε η δεύτερη κάμερα να βρίσκεται παράλληλα ή πλευρικά σε σχέση με την άλλη. Ο σχηματισμός αυτός προσομοιώνει τη διάταξη που είναι τοποθετημένα τα μάτια του ανθρώπου.

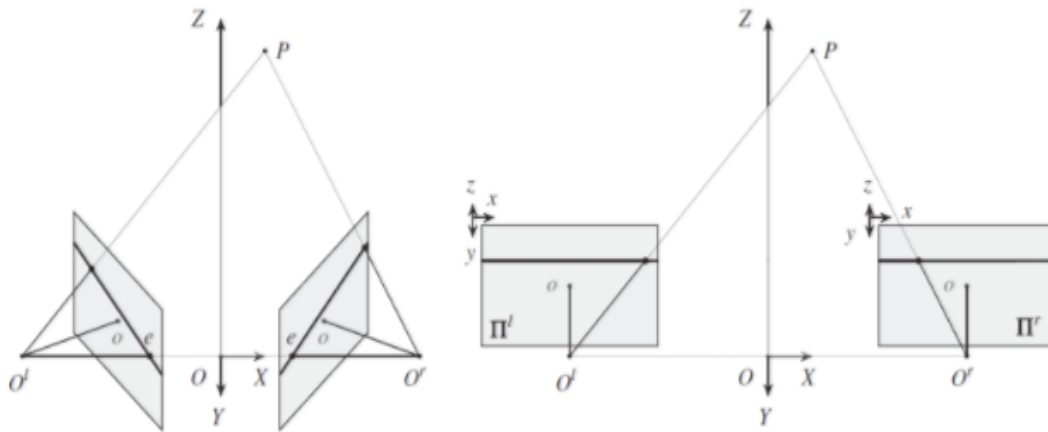
Το στερεοσκοπικό σύστημα χρησιμοποιείται σε πολλές εφαρμογές. Ένα παράδειγμα αποτελεί το ρομποτικό σύστημα της NASA, το οποίο με τη χρήση καμερών είναι ικανό να πλοηγείται με ασφάλεια σε έναν άγνωστο χώρο και να αποφεύγει πιθανά εμπόδια. [3]



Σχήμα 1: Mars Rover [20]

Μέσω της διάταξης αυτής είναι εύκολο να εξαγάγουμε πληροφορίες σχετικά με το βάθος. Αναγνωρίζοντας προβολές του ίδιου σημείου στις δύο εικόνες και παίρνοντας τη διαφορά που προκύπτει από τη θέση στην πρώτη εικόνα και την αντίστοιχη θέση στη δεύτερη, λαμβάνεται ένα στοιχείο σχετικά με το βάθος. Η διαφορά των θέσεων στις εικόνες, στις οποίες προβάλλονται αναγνωρισμένα σημεία της σκηνής, υπάρχει λόγω της διαφορετικής θέσης που έχουν οι κάμερες στον χώρο.

Οι κάμερες ενός στερεοσκοπικού συστήματος μπορεί να είναι είτε παράλληλες είτε γενικότερα συγχλίνουσες, όπως φαίνεται στο παρακάτω σχήμα.



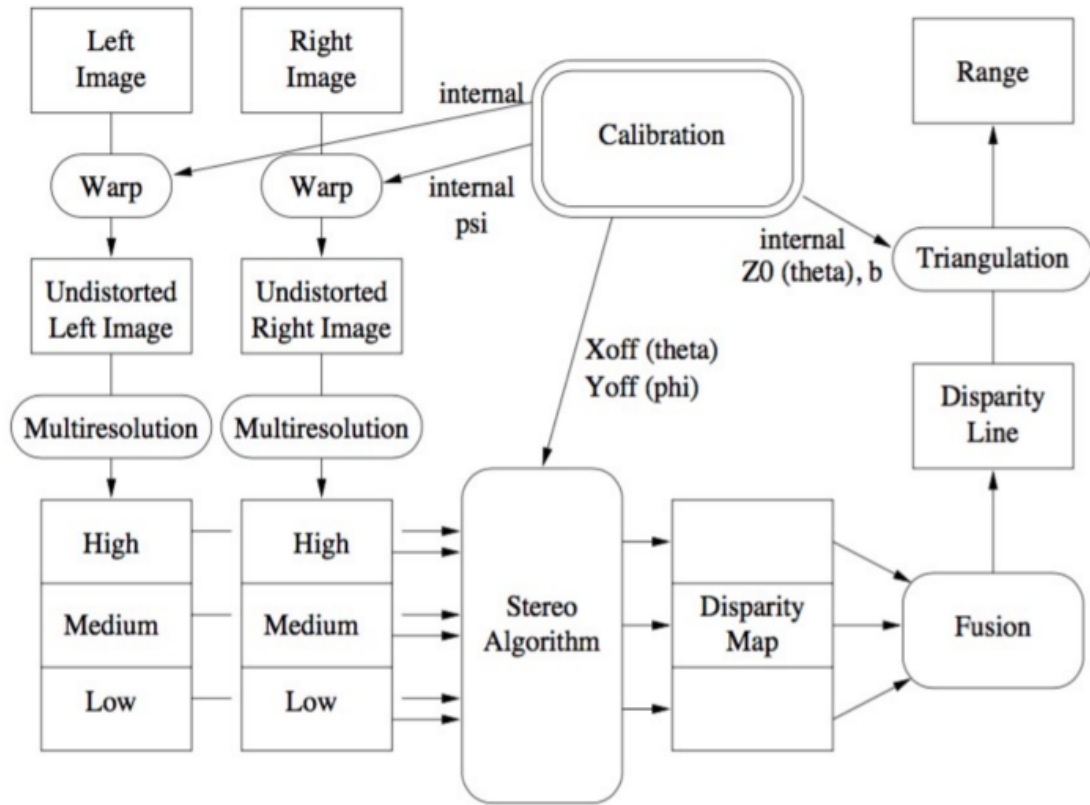
Σχήμα 2: Διατάξεις στερεοσκοπικού συστήματος [18]

## 2.4 Βασικά βήματα για την μοντελοποίηση της 3D σκηνής μέσω του στερεοσκοπικού συστήματος

Για να μοντελοποιηθεί η 3D σκηνή μέσω της στερεοσκοπικής διάταξης χρειάζεται να ακολουθεί μια σειρά από βήματα.

Τα βασικά βήματα είναι τα παρακάτω:

Στο Σχήμα 3 απεικονίζεται αναλυτικά η διαδικασία που ακολουθείται για δημιουργία της σκηνής στον 3D χώρο, μέσω της στερεοσκοπικής όρασης.



Σχήμα 3: Διαδικασία μοντελοποίησης της 3D σκηνής, μέσω του στερεοσκοπικού συστήματος [8]

## 2.5 Βαθμονόμηση κάμερας (Camera calibration)

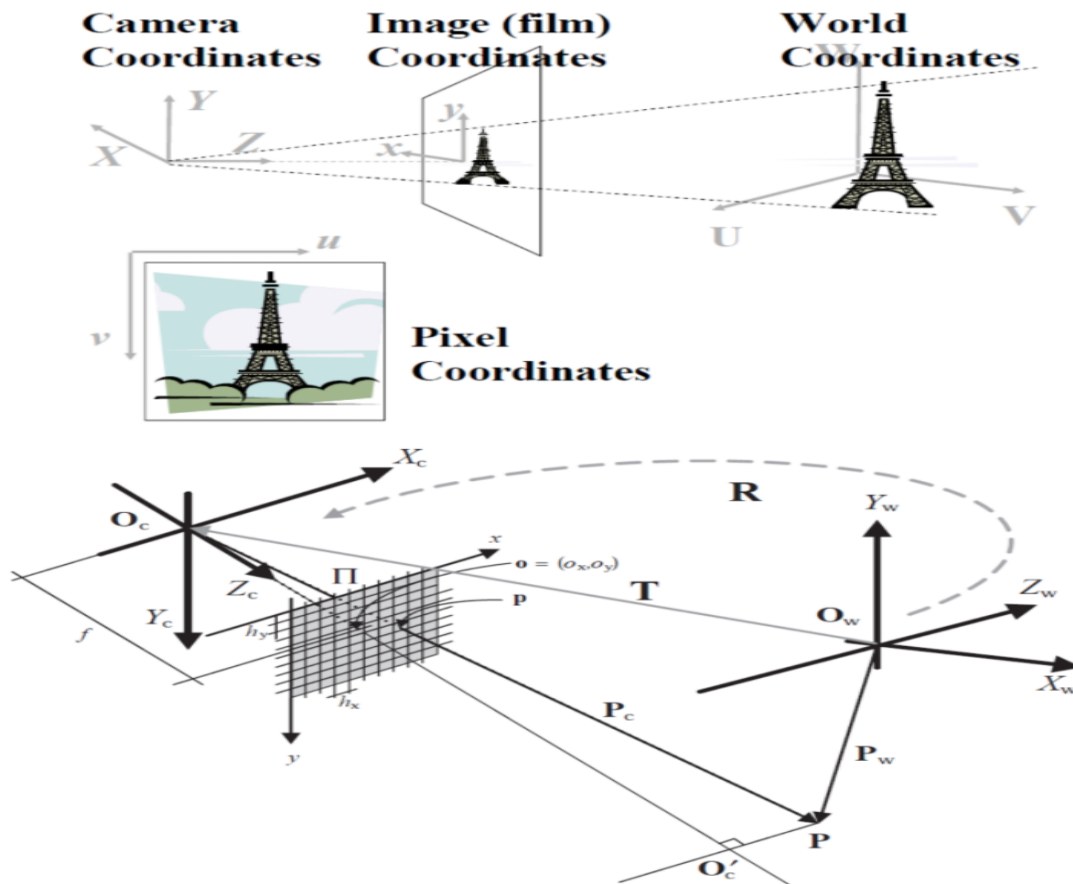
Η βαθμονόμηση κάμερας είναι απαραίτητο βήμα για την απόκτηση πληροφοριών από τις δισδιάστατες εικόνες. Λόγω της σημαντικότητάς της σε διάφορα επιστημονικά πεδία, όπως είναι η υπολογιστική όραση και η φωτογραμμετρία έχει μελετηθεί εκτενώς, ενώ έχουν προταθεί και καινούργιες μέθοδοι βαθμονόμησης. Η βαθμονόμηση κάμερας στο πεδίο της τρισδιάστατης μηχανικής όρασης ορίζεται ως η διαδικασία προσδιορισμού των εσωτερικών γεωμετρικών και οπτικών χαρακτηριστικών μίας κάμερας (εσωτερικές παράμετροι) και/ή την τρισδιάστατη θέση και κατεύθυνση του πλαισίου της κάμερας με ένα συγκεκριμένο γνωστό σύστημα συντεταγμένων (εξωτερικές παράμετροι). Σε αυτό το κεφάλαιο θα αναλυθεί η διαδικασία της βαθμονόμησης.

### 2.5.1 Συστήματα συντεταγμένων

Για να περιγράφουν οι μαθηματικές σχέσεις που προκύπτουν μεταξύ των σημείων του τρισδιάστατου κόσμου και του δισδιάστατου επιπέδου προβολής, πρέπει να οριστούν τα παρακάτω συστήματα συντεταγμένων.

- Σύστημα συντεταγμένων κόσμου ( $U, V, W$ ).
- Σύστημα συντεταγμένων της κάμερας ( $X, Y, Z$ ).
- Σύστημα συντεταγμένων εικόνας ( $x, y$ ), το οποίο είναι η προοπτική προβολή των ( $X, Y, Z$ ) στο επίπεδο εικόνας με εστιακό μήκος  $z = f$ .
- Σύστημα συντεταγμένων εικόνας σε μονάδες pixel ( $u, v$ ).

Οι άξονες  $X$  και  $Y$  της κάμερας είναι παράλληλοι με τους αντίστοιχους άξονες του συστήματος αναφοράς εικόνας. Επίσης, ο  $Z$  άξονας θεωρείται το οπτικό κέντρο της κάμερας. Το σύστημα συντεταγμένων κόσμου συσχετίζεται με το σύστημα συντεταγμένων της κάμερας αν θεωρηθεί μία μετατόπιση που εκφράζεται από τον πίνακα  $T$  και μια περιστροφή που είναι εκφρασμένη από τον πίνακα  $R$ .

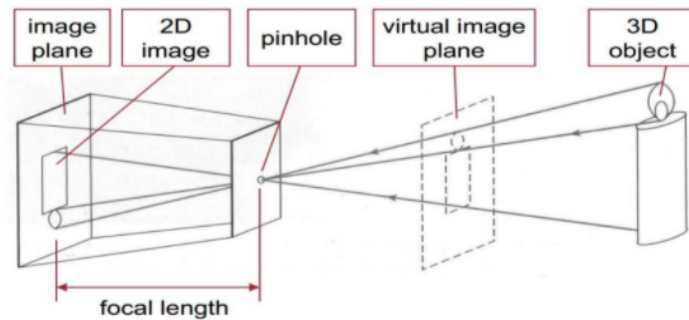


Σχήμα 4: Συστήματα Αναφοράς.

### 2.5.2 Μοντέλο κάμερας (Camera model)

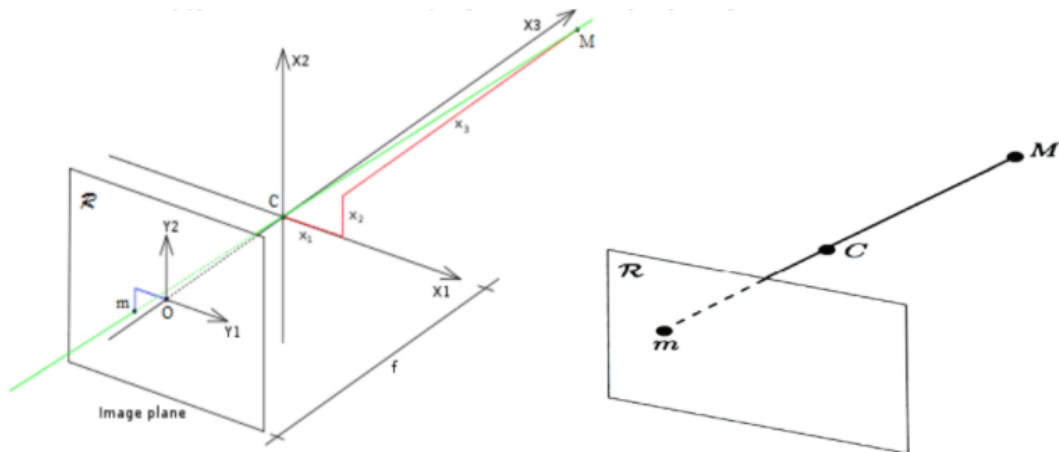
Μια φωτογραφική μηχανή μπορεί να μοντελοποιεί μαθηματικά με αρκετούς τρόπους. Οι τρόποι αυτοί καθορίζονται ανάλογα με τις ιδιότητες, τον βαθμό ακρίβειας, την εφαρμογή για την οποία προορίζεται το μοντέλο και άλλα. Το μοντέλο που χρησιμοποιείται συνήθως για τη μοντελοποίηση τρισδιάστατων αντικειμένων και χώρων είναι το μοντέλο μικρής οπής. [21]

Το μοντέλο μικρής οπής δίνει τη δυνατότητα να περιγράφει η μαθηματική σχέση μεταξύ των συντεταγμένων ενός τρισδιάστατου σημείου και της προβολής του στο επίπεδο της εικόνας, μέσω μίας ιδανικής φωτογραφικής κάμερα μικρής οπής. Το άνοιγμα της κάμερας περιγράφεται ως ένα σημείο, απείρως μικρό, μέσω του οποίου περνάνε όλες οι γραμμές των προβολών των σημείων. Το σημείο αυτό ονομάζεται «μικρή οπή» και ορίζεται ως οπτικό κέντρο της κάμερας για τον τρισδιάστατο χώρο.



Σχήμα 5: Μοντέλο κάμερας «μικρής οπής»

Οι οπτικές ακτίνες ενός αντικείμενου του χώρου, που βρίσκεται μπροστά από την κάμερα, διέρχονται από το οπτικό κέντρο της κάμερας και προσπίπτουν στο επίπεδο της εικόνας. Πιο συγκεκριμένα, το μοντέλο κάμερας «μικρής οπής» μπορεί να εξηγηθεί μαθηματικά βάσει του παρακάτω σχήματος.



Σχήμα 6: Προβολή ενός σημείου

Για να πάρουμε την απεικόνιση  $m$  στο επίπεδο ενός συγκεκριμένου σημείου  $M$ , του τρισδιάστατου χώρου  $P^3$ , παίρνουμε την τομή της οπτικής ακτίνας με το επίπεδο της εικόνας. Η οπτική ακτίνα είναι η νοητή γραμμή, η οποία διέρχεται από τα σημεία  $M$  και  $C$ . Το  $C$  είναι το οπτικό κέντρο από το οποίο διέρχονται οι οπτικές ακτίνες για να απεικονιστούν στο επίπεδο της εικόνας  $R$ . Έτσι το  $C$  είναι η αρχή των συντεταγμένων του τρισδιάστατου χώρου και το  $O$  είναι η αρχή του συστήματος συντεταγμένων της εικόνας, που λέγεται κύριο σημείο. Η απόσταση του  $C$  από το επίπεδο της εικόνας ονομάζεται εστιακή απόσταση (focal length). Η απόσταση αυτή είναι σταθερή και δεν

εξαρτάται από τα  $M$  και  $m$ . Το επίπεδο στο οποίο βρίσκεται το  $C$  ονομάζεται εστιακό επίπεδο και τέλος, η ευθεία που διέρχεται από το σημείο  $C$  και είναι κάθετη στα δυο αυτά επίπεδα ονομάζεται οπτικός άξονας.

Μία πραγματική κάμερα αποτελείται από συστοιχίες φακών, οι οποίες εισάγουν παραμορφώσεις, με αποτέλεσμα τα σημεία να μην είναι συνευθειακά. Όμως υπάρχει τρόπος να τα προσεγγίσουμε τις παραμορφώσεις αυτές μαθηματικά οπότε το μοντέλο μικρής οπής μπορεί να περιγράψει αρκετά καλά την λειτουργία της κάμερας. [15]

### 2.5.3 Προοπτική προβολή (Perspective transform)

Η προοπτική προβολή είναι το γεωμετρικό μοντέλο που μπορεί να χρησιμοποιηθεί για τον σχηματισμό της εικόνας.

Οι συντεταγμένες  $(X, Y, Z)$  ενός αντικειμένου που βρίσκεται στο τρισδιάστατο χώρο του κόσμου μετατρέπονται σε συντεταγμένες εικόνας  $(x, y)$  με την προοπτική προβολή. Πιο συγκεκριμένα, το σημείο  $Q(X, Y, Z)$  στον  $3D$  χώρο απεικονίζεται στο  $2D$  σημείο  $q(x, y, f)$  του επιπέδου απεικόνισης. Υποθέτοντας ότι το κέντρο του συστήματος συντεταγμένων βρίσκεται στο σημείο που είναι η οπή και συμβολίσουμε την εστιακή απόσταση με  $f$ , τότε με ομοιότητες τριγώνων προκύπτουν οι παρακάτω σχέσεις:

$$x = -f \frac{X}{Z}$$

$$y = -f \frac{Y}{Z}$$

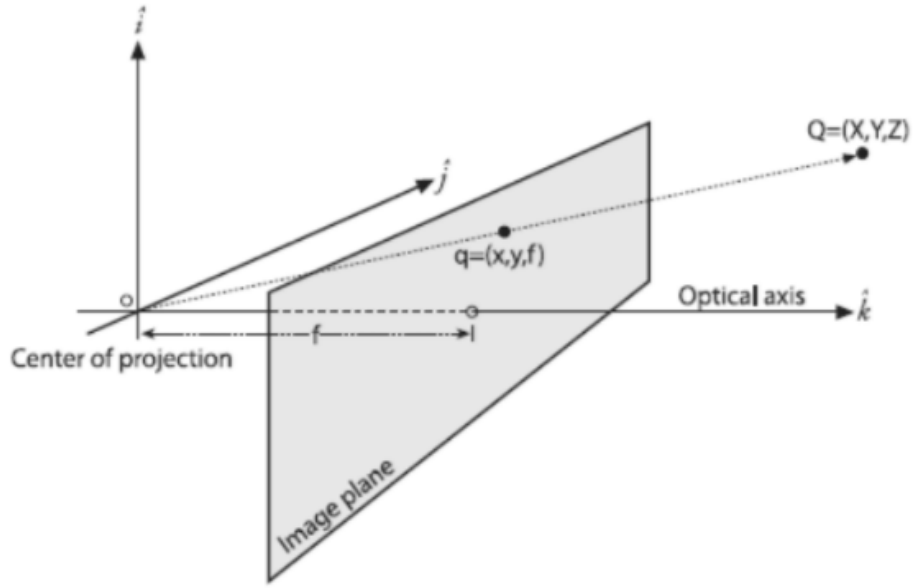
$$z = -f$$

Αν θεωρηθεί ότι το επίπεδο της εικόνας βρίσκεται μπροστά και όχι πίσω από την εστία της κάμερας (Σχήμα 6). Οι παραπάνω σχέσεις γίνονται:

$$x = f \frac{X}{Z}$$

$$y = f \frac{Y}{Z}$$

$$z = f$$



Σχήμα 7: Σύστημα συντεταγμένων με το επίπεδο εικόνας να είναι μπροστά από την εστία της κάμερας

Σε ομογενείς συντεταγμένες ο μετασχηματισμός γράφεται:

$$\tilde{q} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, \tilde{Q} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{fX}{Z} \\ \frac{fY}{Z} \\ 1 \end{bmatrix} = \begin{bmatrix} fX \\ fY \\ Z \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

$$P = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

άρα η σχέση που προκύπτει είναι:

$$\tilde{q} = P\tilde{Q}$$

Ο πίνακας  $P$  ονομάζεται προβολικός πίνακας της κάμερας (projection matrix) περιέχει τις πληροφορίες που χρειαζόμαστε για να προσδιορίσουμε την προβολή οποιουδήποτε σημείου του χώρου στην εικόνα, μέσω της κάμερας.

Για να καθοριστεί η σχέση μεταξύ των συντεταγμένων ενός σημείου του τρισδιάστατου χώρου και των συντεταγμένων της προβολής του στην εικόνα, πρέπει να υπολογιστούν οι παράμετροι της κάμερας. Οι παράμετροι αυτές χωρίζονται σε **ενδογενείς** και **εξωγενείς**. Σε μια πραγματική κάμερα δημιουργούνται παραμορφώσεις, λόγω των ατελειών που υπάρχουν στους φακούς της κάμερας, αυτές οι παράμετροι ονομάζονται ενδογενείς. Οι εξωγενείς παράμετροι εξαρτώνται άμεσα από τη θέση και τον προσανατολισμό της κάμερας σε κάθε λήψη της εικόνας. Η διαδικασία υπολογισμού των παραμέτρων αυτών ονομάζεται **βαθμονόμηση** της κάμερας. [15]

#### 2.5.4 Ενδογενείς παράμετροι (Intrinsic parameters)

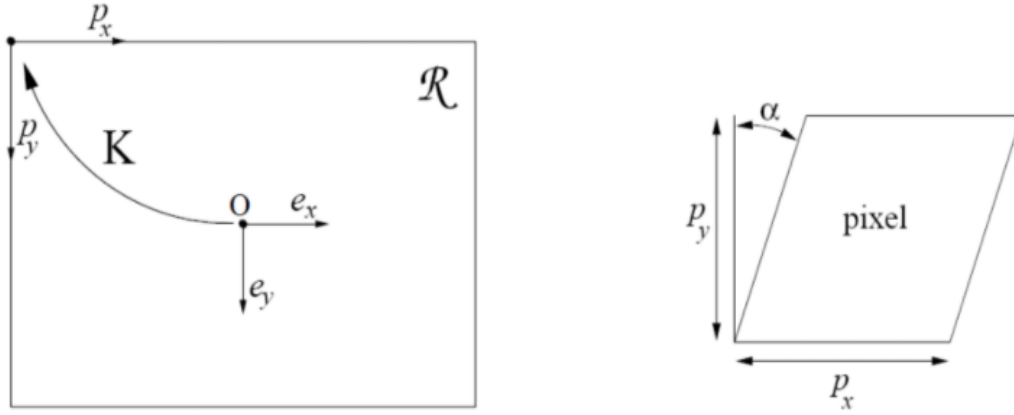
Σε μια πραγματική κάμερα, είναι πολύ πιθανόν οι συντεταγμένες της εικόνας να μην αντιστοιχίζονται με τις συντεταγμένες του εστιακού επιπέδου. Επιπλέον, οι άξονες της εικόνας μπορεί να μην είναι απολύτως ορθογώνιοι. Τα παραπάνω οφείλονται στους αισθητήρες της κάμερας αν δεν είναι με ακρίβεια τοποθετημένοι σε απολύτως ορθογώνια διάταξη.

Είναι σημαντικό να προσδιοριστούν οι εσωτερικές παράμετροι της κάμερας, καθώς έτσι δίνεται η δυνατότητα λήψης μιας τρισδιάστατης απεικόνισης των σημείων της σκηνής με μεγαλύτερη ακρίβεια. Οι παραμορφώσεις αυτές μπορούν να μοντελοποιηθούν και να συμπεριληφθούν στο γραμμικό μοντέλο που έχει ήδη αναλυθεί.

Ο μετασχηματισμός για να ληφθούν τις συντεταγμένες  $(u, v)$  της εικόνας από το εστιακό επίπεδο είναι:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{f}{p_x} & \tan(a) \frac{f}{p_y} & o_x \\ 0 & \frac{f}{p_y} & o_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_R \\ v_R \\ 1 \end{bmatrix}$$

οπού  $p_x$  είναι το πλάτος του pixel και  $p_y$  το ύψος του. Το κύριο σημείο είναι το  $O[o_x, o_y, 1]^T$  και  $a$  είναι η λοξή γωνία όπως φαίνεται στο παρακάτω σχήμα.



Σχήμα 8: Από τις συντεταγμένες του επιπέδου εστίασης στις συντεταγμένες της εικόνας [15]

Η παραπάνω εξίσωση μπορεί να απλουστευτεί:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & s & o_x \\ 0 & f_y & o_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_R \\ v_R \\ 1 \end{bmatrix}$$

$$K = \begin{bmatrix} f_x & s & o_x \\ 0 & f_y & o_y \\ 0 & 0 & 1 \end{bmatrix}$$

οπού  $f_x, f_y$  είναι το εστιακό μήκος μετρημένο με το πλάτος και το ύψος των pixels. Το  $s$  είναι ένας παράγοντας για τη λοξή μετατόπιση που οφείλεται στη μη ορθογωνιότητα των pixels. Στις περισσότερες κάμερες τα pixels είναι σχεδόν τέλεια ορθογώνια, οπότε μπορεί να υποτεθεί ότι το  $s$  τείνει στο 0. Ο πίνακας  $K$  ονομάζεται πίνακας βαθμονόμησης της κάμερας. Λαμβάνοντας υπόψη τις καινούργιες σχέσεις ο προβολικός πίνακας  $P$  μπορεί να γραφτεί ως εξής: [15]

$$P = \begin{bmatrix} f_x & 0 & o_x & 0 \\ 0 & f_y & o_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} f_x & 0 & o_x \\ 0 & f_y & o_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = K[I|0]$$

### 2.5.5 Εξωγενείς παράμετροι (Extrinsic parameters)

Στις παραπάνω σχέσεις έγινε η υπόθεση ότι τα σημεία του 3D κόσμου είναι πάνω σε ένα σύστημα συντεταγμένων με επίκεντρο το σημείο εστίασης, δηλαδή τη φωτογραφική μηχανή. Ωστόσο για ένα πιο πλήρες πραγματικό μοντέλο, πρέπει να βρεθεί μια σχέση

προς το παγκόσμιο σύστημα συντεταγμένων και όχι ως προς το σύστημα συντεταγμένων της κάμερας. Τα τρισδιάστατα σημεία μπορεί να είναι μέρος οποιουδήποτε αυθαίρετου συστήματος συντεταγμένων, όμως υπάρχει ένας μοναδικός μετασχηματισμός (περιστροφή και μεταφορά) που αντικατοπτρίζει την κίνηση της σκηνής, μεταξύ των συστημάτων αυτών. Πιο συγκεκριμένα,

$$\tilde{M}' = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \tilde{M}$$

ο πίνακας  $R$  είναι ένας  $3 \times 3$  πίνακας περιστροφής, το  $t$  είναι ένα διάνυσμα  $3 \times 1$  παράλληλης μετατόπισης. Το ομογενές σημείο  $\tilde{M}'$  αποτελεί ένα σημείο στο σύστημα συντεταγμένων της κάμερας, το οποίο είναι ίδιο με το σημείο του παγκοσμίου συστήματος συντεταγμένων.

Με την αντίστροφη κίνηση της σκηνής ισοδυναμεί η κίνηση της κάμερας. Επομένως η παραπάνω σχέση γίνεται:

$$\tilde{M}' = G\tilde{M}$$

$$G = \begin{bmatrix} R^T & -R^T t \\ 0 & 1 \end{bmatrix}$$

οπού  $R$  και  $t$  αποτελούν τους εξωγενείς παραμέτρους της κάμερας. Αυτοί οι παράμετροι καθορίζουν τον προσανατολισμό και τη θέση της κάμερας σε σχέση με το παγκόσμιο σύστημα συντεταγμένων.

Συνοψίζοντας όλα τα παραπάνω, το συνολικό μαθηματικό μοντέλο που περιγράφει μια ιδανική κάμερα, η οποία έχει συγκεκριμένες ενδογενείς παραμέτρους και συγκεκριμένο προσανατολισμό και θέση, διαμορφώνεται ως εξής:

$$\tilde{q} = K[I|0]G\tilde{Q}$$

οπού,

$$P = K[I|0]G$$

ο πίνακας  $P$  είναι ο πίνακας της κάμερας[2].

### 2.5.6 Παραμορφώσεις (Distortions)

Σε μια πραγματική κάμερα, λόγω των ατελειών που υπάρχουν στους φακούς και στην κατασκευή, εισάγονται παραμορφώσεις στις εικόνες. Αυτό συμβαίνει κυρίως για λόγους απλοποίησης της παραγωγικής διαδικασίας, καθώς είναι πολύ πιο εύκολο να κατασκευαστεί ένας σφαιρικός φακός παρά ένας μαθηματικά ιδανικός φακός. Επίσης, είναι δύσκολο να ευθυγραμμιστεί μηχανικά ο φακός και η συσκευή απεικόνισης. Για ένα αξιόπιστο μοντέλο πρέπει να ληφθούν υπόψη οι παραπάνω παραμορφώσεις. Υπάρχουν δυο σημαντικές παραμορφώσεις η **ακτινική** και η **εφαπτωμενική**.

### 2.5.7 Ακτινική παραμόρφωση(Radial distortion)

Λόγω της ατελούς κατασκευής τους, οι φακοί δεν έχουν την απαιτούμενη καμπυλότητα. Με αποτέλεσμα τα σημεία στην εικόνα να αποκλίνουν από τη θέση που έχει προβλεφθεί με βάση το γραμμικό μοντέλο. Λόγω της ακτινικής παραμόρφωσης οι ευθείες γραμμές έχουν υποστεί ένα είδος στρέβλωσης με αποτέλεσμα να φαίνονται σαν καμπύλες. Το φαινόμενο αυτό γίνεται εντονότερο, όσο απομακρυνόμαστε από το κέντρο της εικόνας. Στο Σχήμα 9 γίνεται εύκολα αντιληπτό το φαινόμενο αυτό. Η εικόνα (α) απεικονίζεται



(α)



(β)

Σχήμα 9: Παράδειγμα ακτινικής παραμόρφωσης [15]

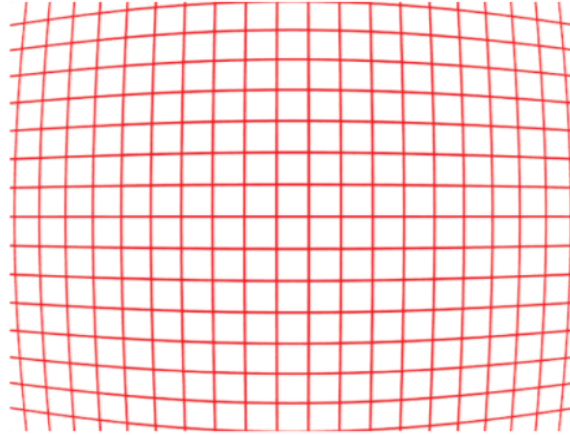
χωρίς παραμόρφωση. Αντιθέτως, η εικόνα (β) έχει υποστεί ακτινική παραμόρφωση ως το προς το κέντρο της εικόνας.

Τα δυο πιο κοινά είδη παραμορφώσεων είναι:

- **Barrel distortion.**
- **Pincushion distortion.**

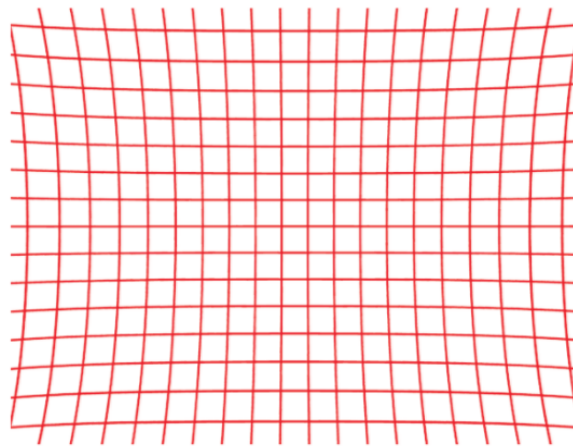
Η παραμόρφωση **barrel distortion** παρατηρείται όταν τα σημεία δεν βρίσκονται στη σωστή θέση με βάση το γραμμικό μοντέλο αλλά έχουν υποστεί θετική μετατόπιση σε σχέση με το κέντρο της εικόνας. Σε αυτή την περίπτωση παρατηρείται θετική ακτινική παραμόρφωση. Αυτό συμβαίνει συχνά στους ευρυγώνιους φακούς, επειδή το οπτικό πεδίο του φακού είναι ευρύτερο από το μέγεθος του αισθητήρα εικόνας και ως εκ τούτου πρέπει να συμπίεστεί για να χωρέσει. Αυτό έχει ως αποτέλεσμα, οι ευθείες γραμμές να

φαίνονται πιο καμπυλωτές προς τα μέσα, ειδικά προς τις άκρες του πλαισίου. Παρακάτω φαίνεται ένα παράδειγμα θετικής ακτινικής παραμόρφωσης.



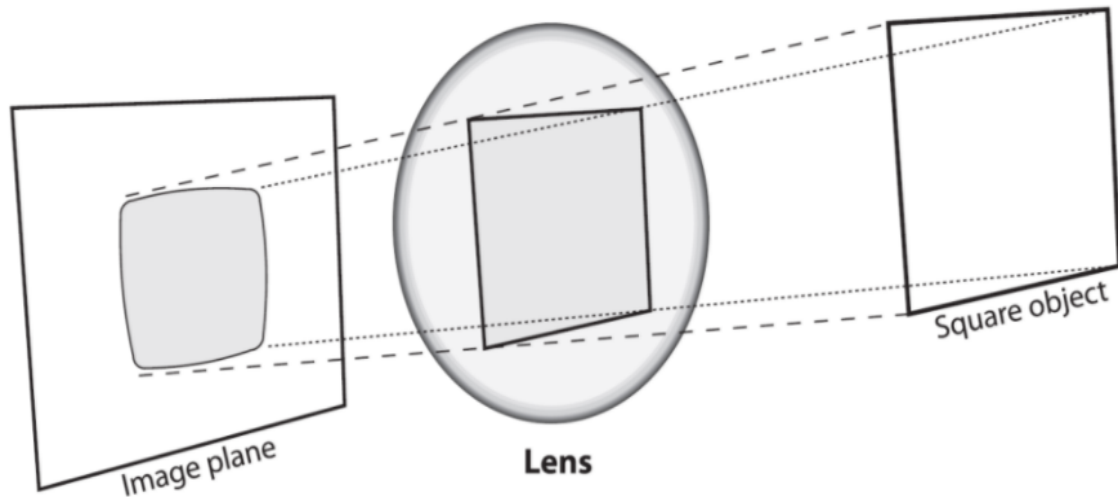
Σχήμα 10: Παράδειγμα θετικής ακτινικής παραμόρφωσης

Η παραμόρφωση **pincushion distortion** είναι παρόμοια με τη barrel, όμως τα σημεία έχουν υποστεί αρνητική μετατόπιση (αρνητική ακτινική μετατόπιση) ως προς το κέντρο της εικόνας. Αυτό του είδους παραμόρφωσης παρατηρείται στους φακούς που χρησιμοποιούνται για ζουμ. Σε αυτή την περίπτωση οι ευθείες φαίνονται πιο καμπυλωτές προς τα έξω.



Σχήμα 11: Παράδειγμα αρνητικής ακτινικής παραμόρφωσης

Η πιο συνηθισμένη ακτινική παραμόρφωση είναι η παραμόρφωση τύπου barrel distortion. Είναι ιδιαίτερα αισθητή στις φθηνές web κάμερες και λιγότερο εμφανής σε φωτογραφικές μηχανές υψηλής τεχνολογίας. Το Σχήμα 12 παρουσιάζει την αίτια που δημιουργεί την ακτινική παραμόρφωση.



Σχήμα 12: Η αίτια της ακτινικής παραμόρφωσης[8]

Μαθηματικά η ακτινική παραμόρφωση μπορεί να διορθωθεί ως εξής:

$$x_{distorted} = x(1 + k_1r^2 + k_2r^4 + k_3r^6)$$

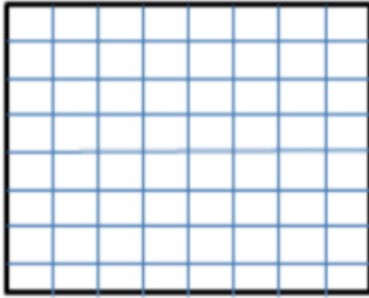
$$y_{distorted} = y(1 + k_1r^2 + k_2r^4 + k_3r^6)$$

όπου,

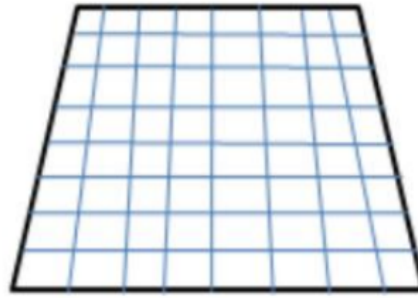
- $x, y$  είναι οι κανονικοποιημένες συντεταγμένες της εικόνας παίρνοντας ως σημείο αναφοράς το κέντρο του φακού,
- $k_1, k_2, k_3$  είναι συντελεστές ακτινικής παραμόρφωσης
- $r^2 = x^2 + y^2$

### 2.5.8 Εφαπτομενική παραμόρφωση (Tangential distortion)

Η εφαπτομενική παραμόρφωση συμβαίνει λόγω ότι δεν επιτυγχάνεται τέλεια παράλληλη ευθυγράμμιση μεταξύ του φακού της κάμερα και του επίπεδου απεικόνισης. Με αποτέλεσμα ορισμένες περιοχές της εικόνας να φαίνονται πιο κοντά ή μακριά ότι είναι στην πραγματικότητα.



***Χωρίς παραμόρφωση***



***Με εφαπτομενική  
παραμόρφωση***

Σχήμα 13: Εφαπτομενική παραμόρφωση

Εφαπτομενική παραμόρφωση μπορεί μαθηματικά να διορθωθεί ως εξής:

$$x_{distorted} = x + [2p_1xy + p_2(r^2 + 2x^2)]$$

$$y_{distorted} = y + [p_1(r^2 + 2y_2) + 2p_2xy]$$

όπου,

- $p_1, p_2$  είναι οι συντελεστές της εφαπτομενικής παραμόρφωσης

Συνοψίζοντας, για τη διόρθωση των παραμορφώσεων χρειάζονται τους πέντε παρακάτω παραμέτρους παραμόρφωσης:

$$Distortion \ coefficients = (k_1 \ k_2 \ p_1 \ p_2 \ k_3)$$

Οι συντελεστές αυτοί εξαρτώνται μονό από την κάμερα και όχι από την προβαλλόμενη σκηνή. Χρειάζεται να υπολογιστούν μια φορά, καθώς συμπεριλαμβάνονται στις εσωτερικές παραμέτρους της κάμερας. Όπως έχει αναφερθεί και παραπάνω, είναι απαραίτητες οι πληροφορίες που αφορούν την κάμερα όπως το εστιακό βάθος( $f_x, f_y$ ) και τα οπτικά κέντρα( $c_x, c_y$ ) της κάμερας. Όλες αυτές τις πληροφορίες δίνονται από τον πίνακα της κάμερας.

$$camera \ matrix = \begin{bmatrix} f_x & 0 & o_x \\ 0 & f_y & o_y \\ 0 & 0 & 1 \end{bmatrix}$$

#### 2.5.9 Μέθοδοι βαθμονόμησης(Calibration methods)

Λόγω της σημαντικότητας της βαθμονόμησης της κάμερας έχουν αναπτυχθεί πολλές τεχνικές και αλγόριθμοι. Οι σημαντικότερες μέθοδοι κατηγοριοποιούνται ως εξής:

- **φωτογραμμετρική βαθμονόμηση** (photogrammetric calibration )
- **αυτό-βαθμονόμηση** (self-calibration)

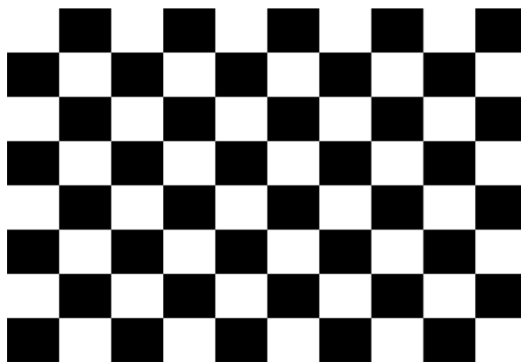
Η **φωτογραμμετρική βαθμονόμηση** γίνεται με την παρατήρηση ενός αντικειμένου, με γνωστή την ακριβή του γεωμετρία στον τρισδιάστατο χώρο. Ανάλογα με τις διαστάσεις του αντικειμένου αυτού οι τεχνικές αυτές χωρίζονται σε τρεις κατηγορίες.

Η πρώτη κατηγορία βαθμονόμησης χρησιμοποιεί ένα **τρειςδιάστατο αντικείμενο** παρατήρησης, του οποίου η γεωμετρία είναι ορισμένη με εξαιρετικά καλή ακρίβεια. Το αντικείμενο αυτό αποτελείται συνήθως από δυο ή τρία επίπεδα κάθετα μεταξύ τους. Συνήθως γίνεται και η χρήση ενός κινούμενου επιπέδου, του οποίου η μετατόπιση είναι γνωστή με αρκετά καλή ακρίβεια. Το πλεονέκτημα αυτής της τεχνικής είναι ότι η βαθμονόμηση μπορεί να γίνει με σχετικά μεγάλη ακρίβεια, ωστόσο όμως λόγω της πολυπλοκότητας της χρειάζεται σχετικά ακριβό εξοπλισμό.



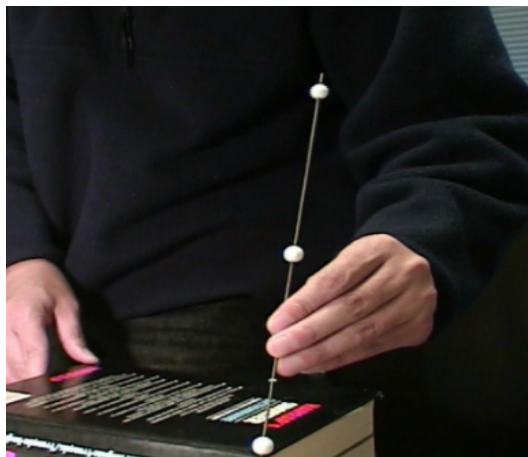
Σχήμα 14: Τρισδιάστατο αντικείμενο παρατήρησης [16]

Η δεύτερη κατηγορία τεχνικών βαθμονόμησης αξιοποιεί την παρατήρηση ενός συγκεκριμένου καλά ορισμένου μοτίβου, το οποίο απεικονίζεται σε ένα επίπεδο. Η παρατήρηση αυτού του **δισδιάστατου αντικειμένου** πρέπει να γίνει από διάφορες άγνωστες κατευθύνσεις. Οι τεχνικές αυτές τείνουν να χρησιμοποιούνται όλο ένα και περισσότερο λόγω της απλότητας που προσφέρουν. Είναι σχετικά εύκολο να κατασκευαστεί το δισδιάστατο μοντέλο παρατήρησης και ο υπολογισμός των παραμέτρων βαθμονόμησης μπορεί να προσεγγιστεί με σχετικά καλή ακρίβεια. Αξίζει να αναφερθεί ότι στην παρούσα διπλωματική χρησιμοποιήθηκε η τεχνική που προτάθηκε από τον Zhengyou Zhang [23], η οποία ανήκει σε αυτή την κατηγορία.



Σχήμα 15: Δισδιάστατο αντικείμενο παρατήρησης

Στην τρίτη κατηγορία οι τεχνικές αυτές χρησιμοποιούν συγγραμμικά σημεία για τον υπολογισμό των παραμέτρων βαθμονόμησης. Πιο συγκεκριμένα, οι τεχνικές βασίζονται στην παρατήρηση μιας γραμμής που κινείται γύρω από ένα σημείο. Τέλος, οι τεχνικές αυτές χρησιμοποιούνται συνήθως στην περίπτωση όπου η παρατήρηση ενός τρισδιάστατου ή δισδιάστατου αντικειμένου, δεν είναι εφικτή.



Σχήμα 16: Μονοδιάστατο αντικείμενο παρατήρησης[16]

Η **αυτό-βαθμονόμηση** ορίζεται ως η διαδικασία υπολογισμού των εσωτερικών παραμέτρων της κάμερας από μη βαθμονομημένες εικόνες. Οι τεχνικές αυτές δεν αξιοποιούν κανένα αντικείμενο βαθμονόμησης. Η βαθμονόμηση επιτυγχάνεται με την κίνηση της κάμερας σε ένα στατικό πεδίο. Επομένως, η λήψη εικόνων από μία κάμερα με αμετάβλητες εσωτερικές παραμέτρους και ο προσδιορισμός αντίστοιχων χαρακτηριστικών του πεδίου, σε τρεις τουλάχιστον εικόνες, αρκούν για τον υπολογισμό των παραμέτρων, που επιτρέπουν την ανακατασκευή τρισδιάστατων δομών με σχετική ακρίβεια. Λόγω του μεγάλου αριθμού των άγνωστων παραμέτρων, τα αποτελέσματα δεν είναι πάντα αρκετά αξιόπιστα. Οι τεχνικές αυτές εφαρμόζονται συνήθως σε εφαρμογές αναγνώρισης και παρακολούθησης, όπου ο βαθμός ακρίβειας δεν είναι σημαντικός.

Αξίζει να αναφερθεί ότι η επιλογή της τεχνικής με την οποία θα πραγματοποιηθεί η βαθμονόμηση της κάμερας πρέπει να γίνεται με βάση τη φύση και τις προδιαγραφές της εφαρμογής. Δεν υπάρχει κάποια τεχνική που να θεωρείται καλύτερη συγκριτικά με τις άλλες για όλες τις εφαρμογές. Στην παρούσα διπλωματική εργασία επιλέχθηκε η τεχνική που προτάθηκε από τον Zhengyou Zhang.

## Η τεχνική βαθμονόμησης του Zhengyou Zhang

Μια από τις πιο αποτελεσματικές τεχνικές βαθμονόμησης θεωρείται η τεχνική του Zhengyou Zhang, καθώς είναι εύκολη στη χρήση και προσφέρει σχετικά καλή ακρίβεια στα αποτελέσματα. Ως μοντέλο βαθμονόμησης χρησιμοποιείται ένα δισδιάστατο ασπρόμαυρο γεωμετρικό μοτίβο. Ο υπολογισμός των παραμέτρων της βαθμονόμησης προκύπτει έπειτα από την παρατήρηση του αντικειμένου βαθμονόμησης από διαφορετικές γωνίες θέασης. Επίσης, η κάμερα και το αντικείμενο μπορούν να κινηθούν, καθώς δεν απαιτείται να είναι γνωστή η σχετική τους κίνηση. [23]

Η διαδικασία βαθμονόμησης κάμερας με την μέθοδο Zhang είναι η εξής:

1. Κατασκευή αντικειμένου παρατήρησης.
2. Λήψη φωτογραφιών από διαφορετικές γωνίες λήψης
3. Εντοπισμός των χαρακτηριστικών σημείου το σχεδίου
4. Υπολογισμός των ενδογενών και εξωγενών παραμέτρων της κάμερας
5. Υπολογισμός των συντελεστών παραμορφώσεων
6. Βελτίωση των παραμέτρων αυτών

### 2.5.10 Ακρίβεια (Accuracy)

Η ακρίβεια μέτρησης αναφέρεται στον βαθμό εγγύτητας που έχουν τα μετρούμενα αποτελέσματα με την πραγματική τιμή. Επίσης, η ακρίβεια δεν είναι μια αριθμητική ποσότητα αλλά ένα χαρακτηριστικό μια μέτρησης. Μια μέτρηση θεωρείται πιο ακριβής όταν έχει ως αποτέλεσμα το μικρότερο σφάλμα μέτρησης.

Όσον αφορά τη βαθμονόμηση για να εξασφαλιστεί μεγάλη ακρίβεια πρέπει να πληρούνται ορισμένες απαιτήσεις. Για παράδειγμα, πρέπει να είναι πολύ καλά κατασκευασμένο το αντικείμενο βαθμονόμησης και με ακρίβεια ορισμένο. Για την καλύτερη εκτίμηση της παραμόρφωσης της κάμερας, το αντικείμενο της βαθμονόμησης πρέπει να εμφανίζεται και στις τέσσερις γωνίες της εικόνας.

Για να εκτιμηθεί κατά ποσό ακριβής είναι η βαθμονόμηση της κάμερας χρειάζεται να υπολογιστεί το σφάλμα **επαναπροβολής (reprojection error)**. Ως τυπικό μέτρο της ακρίβειας της βαθμονόμησης είναι το σφάλμα επαναπροβολής. Το σφάλμα επαναπροβολής υπολογίζεται δίνοντας τις εσωτερικές και εξωτερικές παραμέτρους μια όψης και προβάλλοντας τα γνωστά σημεία του αντικειμένου στην οθόνη. Στη συνέχεια υπολογίζεται το πραγματικό μοντέλο, με βάση του οποίου υπολογίζεται το σφάλμα. Πιο συγκεκριμένα, το σφάλμα επαναπροβολής είναι διάφορα που προκύπτει μεταξύ των σημείων που βρέθηκαν σε μια εικόνα και των ίδιων των σημείων που λήφθηκαν από τον υπολογισμό του πραγματικού μοντέλου.

Οπότε, το σφάλμα των ελάχιστων τετράγωνων προβολής (root mean square (RMS) projection error) μεταξύ των πραγματικών εικονοστοιχείων  $(x_i, y_i)$  και των προβολομένων σημείων  $(x_i^{proj}, y_i^{proj})$ .

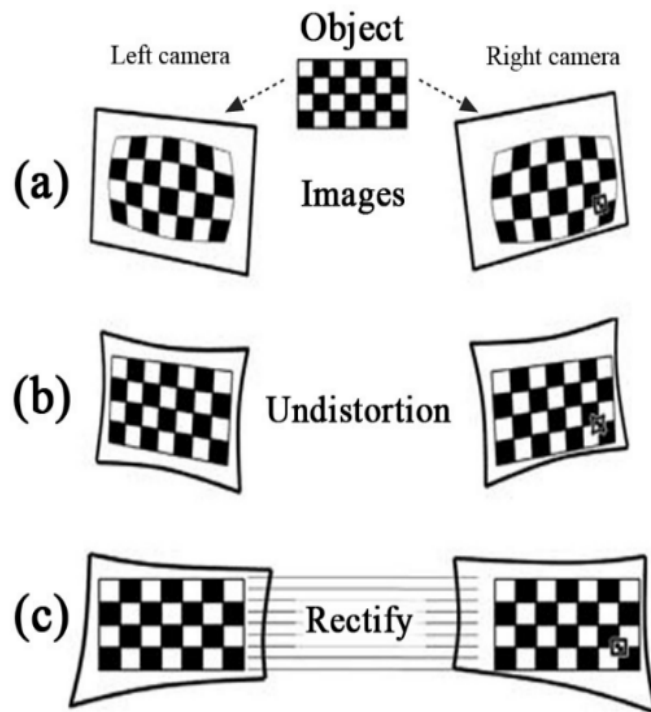
$$E_{reprojection} = \sqrt{\frac{1}{n_{features}} \sum_{i=0}^{n_{features}} ((x_i - x_i^{proj})^2 + (y_i - y_i^{proj})^2)}$$

Το σφάλμα επαναπροβολής μπορεί να χρησιμοποιηθεί για την εκτίμηση της ακρίβειας τόσο για μια κανονική κάμερα όσο και για τη στερεοσκοπική κάμερα. Για πιο ακριβή αποτελέσματα, κρίνεται καλύτερο να υπολογιστούν τα σφάλματα για την κάθε κάμερα ξεχωριστά και στη συνέχεια υπολογιστεί το σφάλμα για τη στερεοσκοπική κάμερα.

## 2.6 Ανόρθωση εικόνας (Image rectification)

Στο κεφάλαιο αυτό θα αναλυθεί η διαδικασία της στερεοσκοπικής ανόρθωσης τόσο θεωρητικά, όσο και σε επίπεδο υλοποίησης.

Η στερεοσκοπική ανόρθωση εικόνων αποτελεί ένα σημαντικό στοιχείο στις εφαρμογές της μηχανικής όρασης. Χρησιμοποιώντας την επιπολική γεωμετρία στις δυο εικόνες και την στερεοσκοπική ανόρθωση τους, μειώνεται σημαντικά το κόστος αναζήτησης στους αλγόριθμους συσχέτισης, καθώς η αναζήτηση γίνεται σε ίδιες οριζόντιες γραμμές, οπότε η αντιστοιχία είναι ένα προς ένα. Πιο συγκεκριμένα, οι στερεοσκοπική μέθοδοι διόρθωσης προσομοιώνουν τις περιστροφές των καμερών για τη δημιουργία επίπεδων εικόνων που είναι παράλληλα στην με τη γραμμή βάσης. Στο Σχήμα 17 γίνεται εύκολα κατανοητή η σημαντικότητα της στερεοσκοπικής ανόρθωσης.

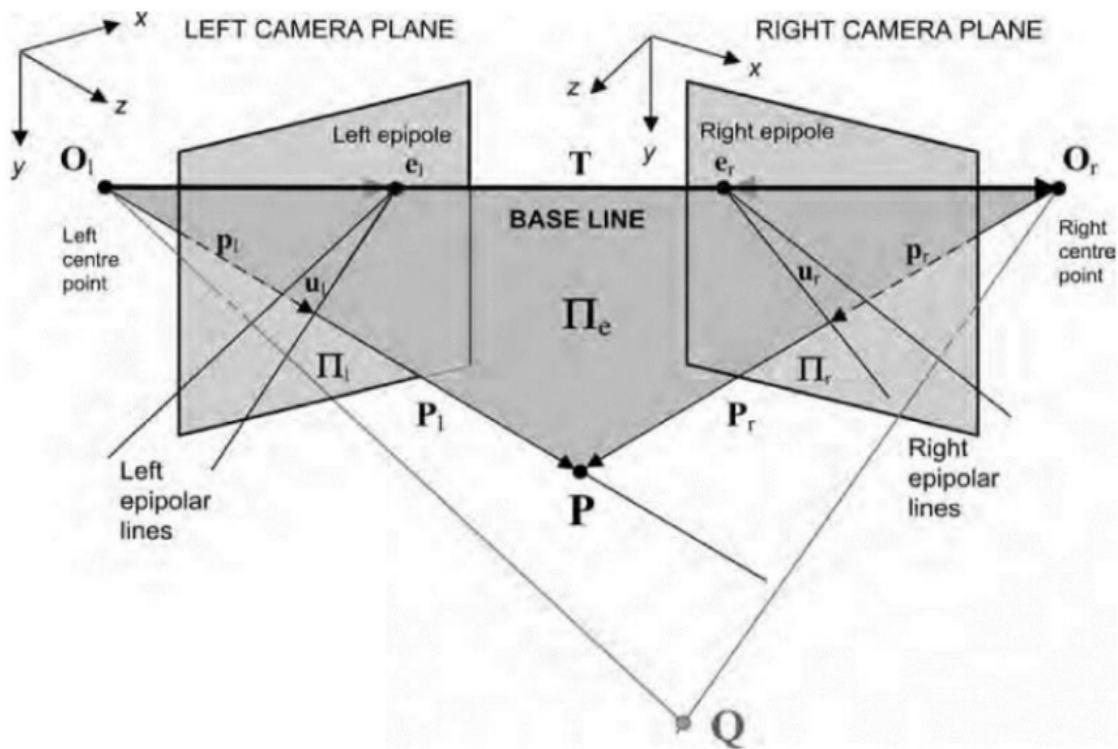


Σχήμα 17: Στερεοσκοπική ανόρθωση εικόνων. [8]

### 2.6.1 Επιπολική γεωμετρία (Epipolar geometry)

Η επιπολική γεωμετρία θεωρείται υποσύνολο της προβολικής γεωμετρίας και παίζει καθοριστικό ρόλο στη στερεοσκοπική όραση. Η επιπολική γεωμετρία αποτελεί την εσωτερική προβολική γεωμετρία μεταξύ δυο εικόνων μιας σκηνής.

Εξαρτάται άμεσα από τη σχετική θέση και τις εσωτερικές παραμέτρους των καμερών. Επίσης, είναι ανεξάρτητη από τη δομή και τον χώρο της σκηνής. Χρησιμοποιώντας το μαθηματικό μοντέλο της μικρής οπής για το ζεύγος των καμερών εξάγει σημαντικές σχέσεις και περιορισμούς ανάμεσα στα σημεία των δυο εικόνων που απεικονίζουν το ίδιο σημείο της σκηνής. Η προβολική γεωμετρία ενός ζεύγους καμερών εκφράζει τη θεμελιώδη σχέση μεταξύ οποιωνδήποτε δυο αντίστοιχων σημείων στα δυο επίπεδα της σκηνής. Στη συνέχεια θα αναλυθεί η επιπολική γεωμετρία με βάση το παρακάτω σχήμα.



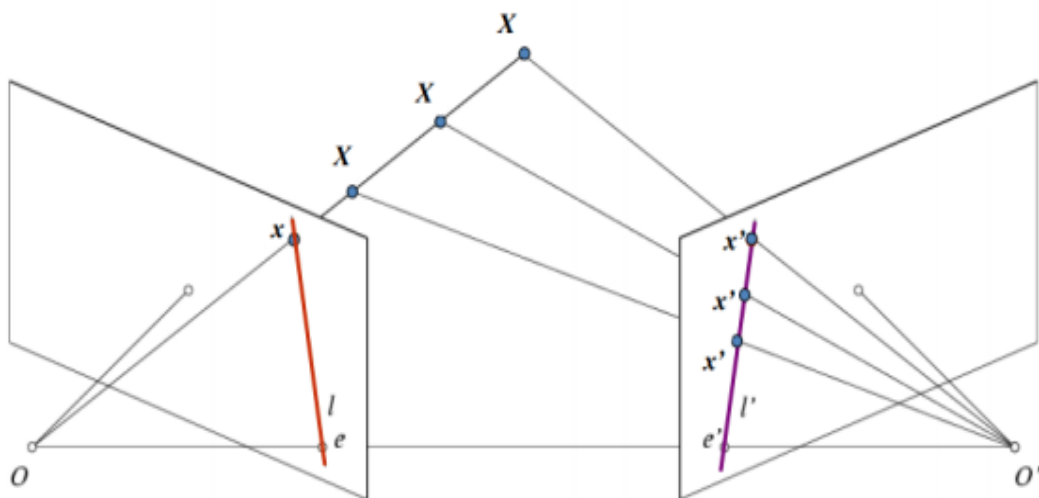
Σχήμα 18: Επιπολική γεωμετρία[8]

Η επιπολική γεωμετρία βασίζεται σε δυο pin-hole κάμερες οι οποίες αποτελούν τα επίπεδα προβολής. Πιο συγκεκριμένα, η αριστερή κάμερα αποτελεί το επίπεδο προβολής  $\Pi_l$  με κέντρο προβολής  $O_l$  και με σύστημα συντεταγμένων  $(x_l, y_l, z_l)$ , ενώ η δεξιά κάμερα αποτελεί το επίπεδο προβολής  $\Pi_r$  με κέντρο προβολής  $O_r$  και με σύστημα συντεταγμένων  $(x_r, y_r, z_r)$ . Η ευθεία που διέρχεται από το κέντρο προβολής  $O$  και είναι κάθετη στο

επίπεδο  $\Pi$ , το σημείο τομής αποτελεί το κύριο σημείο. Η απόσταση μεταξύ του κυρίου σημείου και του κέντρου προβολής θεωρείται το εστιακό μήκος. Η ευθεία που ενώνει τα δυο κέντρα προβολής ονομάζεται βασική γραμμή (base line). Τα σημεία όπου η βασική γραμμή τέμνει το επίπεδο  $\Pi$  καλούνται επιπολικά σημεία και συμβολίζονται με  $e$ . Οπότε τα επιπολικά σημεία για τα επίπεδα προβολής των καμερών είναι  $e_r$  και  $e_l$  αντίστοιχα.

Οι προβολές ενός γνωστού σημείου  $P$  του χώρου στα επίπεδα προβολής, για την αριστερή κάμερα είναι το σημείο  $p_l(x_l, y_l, z_l)$ , ενώ αντίστοιχα για τη δεξιά κάμερα είναι  $p_r(x_r, y_r, z_r)$ . Επίσης, το επίπεδο που σχηματίζεται από τα σημεία  $P$  και τα κέντρα προβολής  $O_r, O_l$  ονομάζεται επιπολικό επίπεδο  $\Pi_e$ . Αυτό το επίπεδο τέμνει τα άλλα επίπεδα προβολής και έτσι δημιουργούνται οι επιπολικές γραμμές.

Για να γίνει πιο κατανοητός ο ρόλος και η σημαντικότητα των επιπολικών γραμμών, αν θεωρηθεί ότι μονό η αριστερή κάμερα και αναλυθεί η προβολή του σημείου  $P$ . Τότε ορίζεται η ευθεία που περνάει από τα σημεία  $O_l, p_l$ . Το σημείο προβολής του παραπάνω σημείου είναι σημείο προβολής όλων των άλλων σημείων που βρίσκονται πάνω στην ίδια ευθεία με το σημείο  $P$ . Όπως γίνεται εύκολα αντιληπτό το σημείο  $P$  μπορεί να είναι οποιοδήποτε σημείο πάνω στην ευθεία  $O_l, p_l$ . Επομένως, είναι αδύνατων να βρεθεί η ακριβής του θέση στον 3D χώρο με μια μονό εικόνα. Οπότε για να εντοπιστεί η θέση του σημείου στον χώρο χρειάζεται άλλο ένα σημείο προβολής της εικόνας από μια άλλη θέση. Αυτό το σημείο μπορεί να προέρχεται από τη δεξιά κάμερα, δηλαδή από το επίπεδο  $\Pi_r$ . Πιο συγκεκριμένα, το σημείο αυτό είναι το  $p_r$ . Οπότε το σημείο  $p_r$  και το κέντρο  $O_r$  φτιάχνουν μια δεύτερη ευθεία. Με τη βοήθεια αυτής της ευθείας θα προσδιοριστεί η ακριβής θέση του σημείου  $P$  στον χώρο. Η ευθεία διέρχεται από το κέντρο  $O_r$  και τέμνει την ευθεία που διέρχεται από τα σημεία  $O_l$  και  $p_l$  στο σημείο  $P$ . Επιπρόσθετα, το σημείο τομής κάθε ευθείας από τις παραπάνω με τα αντίστοιχα επίπεδα προβολής βρίσκεται στις επιπολικές γραμμές. Στο παρακάτω σχήμα είναι αρκετά κατανοητή η σημαντικότητα των επιπολικών γραμμών.



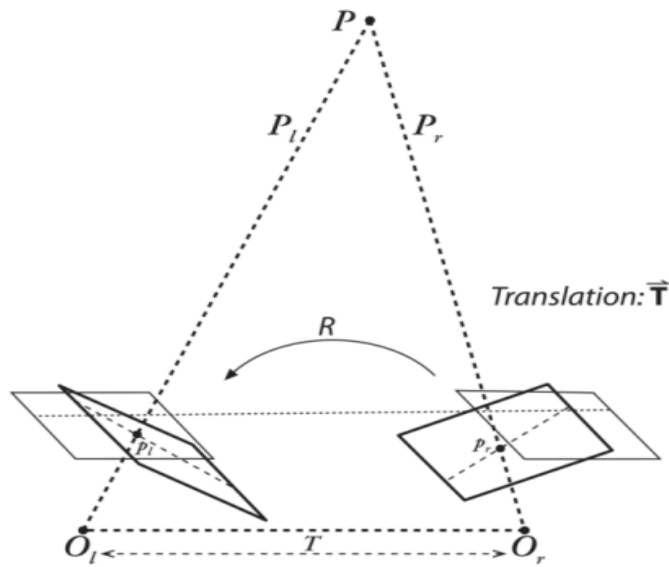
Σχήμα 19: Επιπολικές γραμμές [13]

Κάθε σημείο που απεικονίζεται στην μια εικόνα, το ίδιο σημείο πρέπει να παρατηρηθεί και στην άλλη εικόνα και να βρίσκεται πάνω σε μια γνωστή επιπολική γραμμή. Πιο συγκεκριμένα, όλα τα σημεία που είναι πάνω στην ευθεία που σχηματίζεται από ένα τρισδιάστατο σημείο και με το κέντρο προβολής του, προβάλλονται πάνω στο ίδιο σημείο στο επίπεδο απεικόνισης. Όλα τα σημεία που βρίσκονται στην παραπάνω ευθεία, σχηματίζουν μια επιπολική γραμμή στο άλλο επίπεδο απεικόνισης, δηλαδή την άλλη κάμερα. Μελετώντας το Σχήμα 26 μπορεί να γίνει πιο κατανοητό το τι συμβαίνει, έστω το 3D σημείο  $X$  στο χώρο και το σημείο  $O$  αποτελεί το κέντρο προβολής του, έτσι σχηματίζεται η ευθεία που διέρχεται από αυτά τα σημεία ( $OX$ ). Επίσης, προβάλλονται πάνω στο επίπεδο απεικόνισης στο σημείο  $x$ . Όλα αυτά τα σημεία που βρίσκονται στην ευθεία αυτή και με το κέντρο προβολής τους  $O'$  προβάλλονται στο επίπεδο απεικόνισης πάνω στη μωβ γραμμή, δηλαδή σχηματίζουν την επιπολική γραμμή. Το γεγονός αυτό ονομάζεται επιπολικός περιορισμός. Ο επιπολικός περιορισμός βελτιώνει πολύ το κόστος αντιστοιχίας σημείων και ο λόγος είναι ότι δεν χρειάζεται να γίνει αναζήτηση σε ολόκληρο το επίπεδο της εικόνας, αλλά μόνο σε ένα συγκεκριμένο κομμάτι. Με βάση το παραπάνω σχήμα, για το σημείο  $x$  οι πιθανές αντιστοιχίες του πρέπει να βρίσκονται πάνω στη μωβ ευθεία, ενώ οι αντιστοιχίες για το σημείο  $x'$  μπορεί να είναι στην κόκκινη ευθεία.

### 2.6.2 Στερεοσκοπική ανόρθωση (Stereo rectification)

Στην προηγούμενη παράγραφο, αναλύθηκε μαθηματικά η επιπολική γεωμετρία, στην παράγραφο αυτή θα γίνει εφαρμογή της επιπολικής γεωμετρίας στο στερεοσκοπικό σύστημα. Αρχικά, κάθε κάμερα που αποτελεί το στερεοσκοπικό ζεύγος θεωρείται και ένα ξεχωριστό σύστημα συντεταγμένων με τα κέντρα τους να είναι όμοια με τα κέντρα προβολής της κάμερας. Επίσης, έστω το 3D σημείο  $P$  στο χώρο, τότε το διάνυσμα  $P_l = [X_l, Y_l, Z_l]$  απεικονίζει το σημείο αυτό στο σύστημα συντεταγμένων της αριστερής κάμερας, αντίστοιχα το διάνυσμα  $P_r = [X_r, Y_r, Z_r]$  αντιπροσωπεύει το σημείο αυτό για το σύστημα της δεξιάς κάμερας. Επιπλέον, όπως έχει αναλυθεί στο προηγούμενο κεφάλαιο σχετικά με το μαθηματικό μοντέλο της κάμερας, το διάνυσμα  $p_l = [x_l, y_l, z_l]$  αποτελεί την προβολή του σημείου  $P$  για την αριστερή κάμερα, ενώ το διάνυσμα  $p_r = [x_r, y_r, z_r]$  είναι η προβολή του για τη δεξιά κάμερα. Όμως, το  $z_l$  ισούται με το εστιακό μήκος της αριστερής κάμερας, αντίστοιχα το  $z_r$  και για τη δεξιά κάμερα.

Είναι σχετικά εύκολο για γίνει η αντιστοιχία των σημείων όταν τα επίπεδα απεικόνισης είναι τέλεια ευθυγραμμισμένα. Σε ένα πραγματικό στερεοσκοπικό σύστημα αυτό είναι πολύ σπάνιο να εφαρμοστεί, διότι είναι πολύ δύσκολο να φτιαχτεί ένα τέλειο μαθηματικά και με ακρίβεια ορισμένο στερεοσκοπικό σύστημα. Θεωρητικά σε ένα τέλειο κατασκευασμένο στερεοσκοπικό σύστημα, όπου οι εικόνες θα ήταν συνεπίπεδες και οι προβολές του 3D σημείου θα είναι οριζόντια μετατοπισμένες. Πιο συγκεκριμένα, οι επιπολικές γραμμές που σχηματίζονται θα είναι παράλληλες με την βασική οριζόντια γραμμή (base line). Όπως αναφέρθηκε, στην πραγματικότητα είναι σχεδόν αδύνατον να κατασκευαστεί μια τέλεια στερεοσκοπική κάμερα όπου οι φακοί θα είναι τέλεια τοποθετημένοι και ευθυγραμμισμένη μεταξύ τους. Γι' αυτό απαραίτητο κρίνεται πριν εφαρμοστεί η διαδικασία της αντιστοίχισης των σημείων, να γίνει στερεοσκοπική ανόρθωση στις εικόνες. Με αυτόν τον τρόπο τα αποτελέσματα θα είναι πολύ καλύτερα και το κόστος της αναζήτησης θα βελτιωθεί αισθητά, καθώς είναι σχεδόν σίγουρο ότι σε εκείνη τη γραμμή βρίσκονται τα σημεία στις εικόνες.



Σχήμα 20: Στερεοσκοπική ανόρθωση [8]

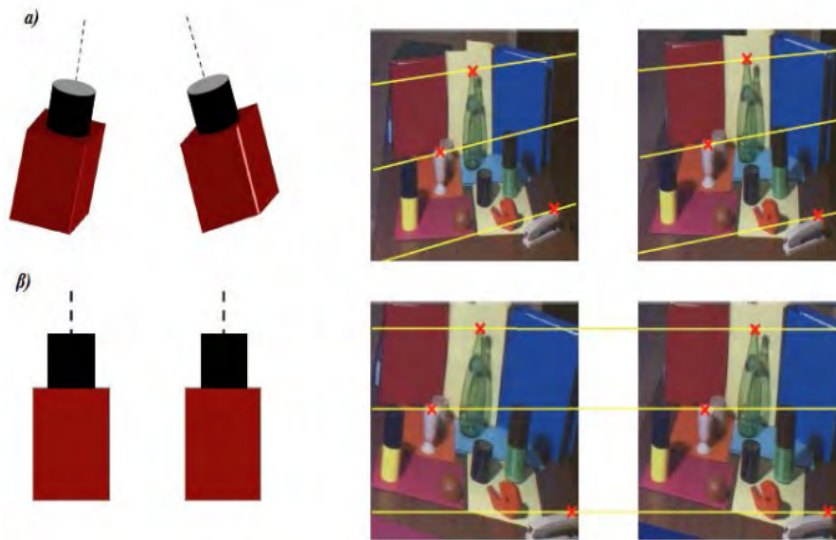
Στο Σχήμα 27, είναι εμφανές ότι τα επίπεδα απεικόνισης θα περιστραφούν και θα μετατοπιστούν κατά τέτοιο τρόπο, ώστε να είναι παράλληλα και ευθυγραμμισμένα μεταξύ τους. Ουσιαστικά, οι επιπολικές γραμμές που σχηματίζονται μετακινούνται προς το άπειρο, αφού δεν είναι δυνατόν το οπτικό πεδίο τους να συμπεριλαμβάνει τα οπτικά τους κέντρα.

Γενικά, ως ανόρθωση εικόνας ορίζεται η διαδικασία κατά την οποία τα ζευγάρια των επιπολικών γραμμών γίνονται παράλληλα και συνευθειακά με την οριζόντια βασική γραμμή. Επίσης, η ανόρθωση των εικόνων είναι ανεξάρτητη από τη γεωμετρία της σκηνής, καθώς μονό τα επίπεδα των εικόνων αλλάζουν θέση, πιο συγκεκριμένα περιστρέφονται γύρω από τους οπτικούς άξονες των καμερών.

Πρέπει να αναφερθεί ότι πολύ σημαντικό βήμα πριν τη διόρθωση της στερεοσκοπικής εικόνας, είναι η βαθμονόμηση των καμερών, όπως έχει περιγράψει στα προηγούμενο κεφάλαιο. Καθώς οι παραμορφώσεις που εισάγονται από τις ατέλειες των καμερών παίζουν σημαντικό ρόλο γιατί τα σημεία μπορεί να έχουν μετακινηθεί από την αναμενόμενη θέση τους με αποτέλεσμα να μην μπορούν να αντιστοιχίσουν σωστά. Η σωστή βαθμονόμηση της στερεοσκοπικής κάμερας κρίνεται απαραίτητη. Για τη βαθμονόμηση των καμερών, μπορεί να χρησιμοποιηθεί η εφαρμογή που έχει αναπτυχθεί στα πλαίσια της διπλωματικής εργασίας. Έχοντάς υπολογισμένες τις εσωτερικές παραμέτρους και τους πίνακες περιστροφής και μεταφοράς των καμερών, είναι δυνατόν να εφαρμοστεί η ανόρθωση της στερεοσκοπικής εικόνας.

Έχουν αναπτυχθεί αρκετοί αλγόριθμοι για την ανόρθωση των εικόνων στο πεδίο της μηχανικής όρασης. Ένας από τους πιο γνωστούς αλγορίθμους είναι ο αλγόριθμος του

Bouguet, οποίος έχει ως στόχο την ελαχιστοποίηση των παραμορφώσεων και την αύξηση του οπτικού πεδίου χρησιμοποιώντας τις παραμέτρους μεταφοράς και περιστροφής. Στο Σχήμα 21 φαίνονται τα αποτελέσματα της στερεοσκοπικής ανόρθωσης με την εφαρμογή του, στην εικόνα α του Σχήματος 21 απεικονίζονται πως είναι κανονικά οι εικόνες και οι επιπολικές γραμμές, ενώ στην εικόνα β του Σχήματος 21 παρατηρείται ότι οι κάμερες έχουν γίνει παράλληλες και οι επιπολικές γραμμές έχουν ευθυγραμμιστεί.



Σχήμα 21: Στερεοσκοπική ανόρθωση αποτελέσματα [8]

### 2.6.3 Αλγόριθμος Bouguet (Bouguet's algorithm)

Ο αλγόριθμος του αναπτύχθηκε από τον Bouguet έχει ως κύριο στόχο την ελαχιστοποίηση των παραμορφώσεων από την επαναπροβολή της εικόνας μέσω της μεταφοράς και της περιστροφής, ενώ παράλληλα στοχεύει και στη μεγιστοποίηση του οπτικού πεδίου.

Για να ελαχιστοποιηθεί η παραμόρφωση επαναπροβολής της εικόνας, ο πίνακας περιστροφής  $R$  που έχει υπολογιστεί κατά τη διαδικασία της βαθμονόμησης. Ο πίνακας περιστροφής είναι ο πίνακας ο οποίος περιστρέφει το επίπεδο εικόνας της δεξιάς κάμερας με το επίπεδο της αριστερής κάμερας. Αυτός ο πίνακας διασπάται σε άλλους δυο πίνακες, τον πίνακα  $r_l$  που αφορά την αριστερή κάμερα και τον πίνακα  $r_r$ , που είναι για τη δεξιά κάμερα. Με αυτόν τον τρόπο κάθε κάμερα περιστρέφεται κατά μίση περιστροφή, έτσι εξασφαλίζεται ότι τα επίπεδα απεικόνισης θα είναι συνεπίπεδα, όχι όμως ευθυγραμμισμένα.

Για να είναι ευθυγραμμισμένες οι επιπολικές γραμμές με την οριζόντια γραμμή, πρέπει να υπολογιστεί ο πίνακας  $R_{rect}$ . Ο πίνακας αυτός θα πάει τα επιπολικά της αριστερής κάμερας στο άπειρο και θα τα ευθυγραμμίσει τις επιπολικές γραμμές οριζόντια. Δημιουργείται ένας πίνακας περιστροφής ξεκινώντας από την κατεύθυνση του ίδιου του επιπόλου  $e_1$  και λαμβάνοντας υπόψη το κύριο σημείο  $c_x, c_y$  ως σημείο αναφοράς της αριστερής ει-  
κόνας, τότε η κατεύθυνση του επιπόλου είναι ακριβώς κατά μήκος του διανύσματος μετατόπισης των κέντρων προβολής των δυο καμερών.

$$e_1 = \frac{T}{||T||}$$

Το διάνυσμα  $e_2$  πρέπει να είναι ορθογώνιο ως προς το  $e_1$ .

$$e_2 = \frac{[-T_y T_x O]^T}{\sqrt{T_x^2 + T_y^2}}$$

Το διάνυσμα  $e_3$  είναι ορθογώνιο ως προς τα άλλα δυο διανύσματα. Για το διάνυσμα  $e_3$  ισχύει

$$e_3 = e_1 \times e_2$$

Έτσι ο πίνακας  $R_{rect}$  που προκύπτει είναι:

$$R_{rect} = \begin{bmatrix} (e_1)^T \\ (e_2)^T \\ (e_3)^T \end{bmatrix}$$

Αυτός ο πίνακας περιστρέφει την αριστερή κάμερα γύρω από το κέντρο προβολής με αποτέλεσμα οι επιπολικές γραμμές να γίνονται οριζόντιες και τα επιπολικά να είναι στο άπειρο. Η ευθυγράμμιση μεταξύ των δυο καμερών επιτυγχάνεται με τον παρακάτω υπολογισμό: [8]

$$R_1 = R_{rect} r_l$$

$$R_2 = R_{rect} r_r$$

## 2.7 Στερεοσκοπική αντιστοίχιση (Stereo matching)

Η βαθμονόμηση των καμερών είναι μια ακριβής και τυπική διαδικασία. Η στερεοσκοπική ανόρθωση είναι η μαθηματική διαδικασία που εφαρμόζεται στις εικόνες για καλύτερα αποτελέσματα. Η πιο δύσκολη διαδικασία για την τρισδιάστατη ανακατασκευή εικόνων είναι η στερεοσκοπική αντιστοίχιση των σημείων.

Ένα από τα βασικά προβλήματα που αντιμετωπίζει η στερεοσκοπική όραση είναι η εύρεση αντιστοιχιών ανάμεσα στις δυο εικόνες. Η διαδικασία αυτή ονομάζεται στερεοσκοπική αντιστοίχιση. Όλα αυτά που έχουν αναλυθεί και παρουσιαστεί έχουν ως προϋπόθεση τη γνώση των δυο ζευγαριών pixels που αντιπροσωπεύουν το ίδιο τρισδιάστατο σημείο της σκηνής.

### 2.7.1 Αλγόριθμοι στερεοσκοπικής αντιστοίχισης (Stereo matching algorithm)

Η στερεοσκοπική αντιστοίχιση είναι η βάση της στερεοσκοπικής όρασης, έχουν αναπτυχθεί αρκετοί αλγόριθμοι, όμως οι μελετητές συνεχίζουν να αναζητούν καλύτερες και βέλτιστες μεθόδους για τον εντοπισμό της σωστής αντιστοίχισης όλων των ζευγαριών των εικονοστοιχείων στις δυο εικόνες.

Οι αλγόριθμοι που έχουν αναπτυχθεί έχουν αρκετές διαφορές μεταξύ τους, όμως έχουν μερικά κοινά στοιχεία. Παρακάτω αναφέρονται μερικά κρίσιμα στοιχεία:

1. Υπολογισμός αντιστοίχισης κόστους (Matching cost computation)
2. Συνάθροιση κόστους (Cost aggregation)
3. Υπολογισμός στερεοσκοπικών διαφορών (Disparity computation)
4. Βελτίωση στερεοσκοπικών διαφορών (Disparity refinement)

Όλοι οι στερεοσκοπικοί αλγόριθμοι απαιτούν ένα κριτήριο κόστους για τη μέτρηση της αντιστοίχισης μεταξύ των δυο pixels. Ο **υπολογισμός του κόστους αντιστοίχισης** είναι το στάδιο στο οποίο οι τιμές των δυο εικονοστοιχείων αντιστοιχούν στο ίδιο σημείο σε μια σκηνή προσδιορίζεται κατά αυτόν τον τρόπο. Ο υπολογισμός του κόστους αντιστοίχισης μπορεί να οριστεί ως η μέθοδος εντοπισμού των τιμών κάθε σημείου μεταξύ της αριστερής και της δεξιάς εικόνας. Το κόστος της αντιστοίχισης υπολογίζεται για κάθε εικονοστοιχείο της εικόνας. Η αντιστοίχιση μπορεί να πραγματοποιηθεί μέσω μια μονοδιάστατης οριζόντιας αναζήτησης, αρκεί οι εικόνες να έχουν υποστεί ακριβή στερεοσκοπική διόρθωση. Οπότε, πολύ σημαντικό παίζει η στερεοσκοπική ανόρθωση στη στερεοσκοπική αντιστοίχιση. Αυτό συμβαίνει, γιατί η αναζήτηση των αντιστοιχιών περιορίζεται μονό σε μια γραμμή και όχι σε όλη την εικόνα, με αποτέλεσμα να μειώνεται αρκετά ο χρόνος και το εύρος αναζήτησης.

Η **συνάθροιση κόστους** είναι το πιο σημαντικό στάδιο, καθώς δείχνει τη γενική απόδοση του αλγορίθμου για τον χάρτη ανισότητας. Ο σκοπός της συνάθροισης κόστους είναι να ελαχιστοποιηθούν οι αβεβαιότητες που προκύπτουν κατά τη διαδικασία της αντιστοίχισης. Η συνάθροιση κόστους είναι αναγκαία γιατί οι πληροφορίες που λαμβάνονται μόνο από ένα pixel δεν αρκούν για τον ακριβή υπολογισμό του κόστους αντιστοίχισης.

Ο **υπολογισμός στερεοσκοπικών διαφορών** διαφέρει ανάλογα με τις προσεγγίσεις των αλγορίθμων. Υπάρχουν δυο κυρίες προσεγγίσεις:

1. τοπική προσέγγιση (local approach )
2. καθολική προσέγγιση (global approach)

Η **τοπική προσέγγιση** λαμβάνει υπόψη ένα μικρό αριθμό γειτονικών pixels γύρω από το εικονοστοιχείο που είναι προς μελέτη. Σε γενικές γραμμές αυτή η μέθοδος είναι πολύ αποτελεσματική, αλλά είναι σχετικά ευαίσθητη στις τοπικές ασάφειες των περιοχών. Για την δημιουργία του χάρτη ανισότητας, η διάφορα για κάθε pixel επιλέγεται ουσιαστικά χρησιμοποιώντας την στρατηγική ότι ο νικητής τα παίρνει όλα (WTA strategy). Μαθηματικά ορίζεται ως εξής:

$$d_p = \arg_{d \in D} \min C'(p, d)$$

Η διάφορα που σχετίζεται με το ελάχιστο συγκεντρωτικό κόστος  $d_p$  σε κάθε pixel είναι και η τιμή που επιλέγεται. Επίσης,  $C'_p$  αντιπροσωπεύει το συνολικό κόστος που προκύπτει μετά το αντίστοιχο υπολογιστικό κόστος και  $D$  υποδηλώνει το σύνολο όλων των επιτρεπόμενων διακριτών ανισοτήτων.

Αντίθετα, για την **καθολική προσέγγιση** γίνονται ορισμένες υποθέσεις σχετικά με το βάθος του πεδίου της σκηνής, που συνήθως εκφράζονται με μια ενεργειακή ελαχιστοποίηση. Γενικά, ο στόχος είναι να βρεθεί η βέλτιστη συνάρτηση εκχώρησης της ενεργειακής ανισότητας  $d = d(x, y)$  η οποία ελαχιστοποιεί

$$E(d) = E_{data} + \beta E_{smooth}(d)$$

οπού  $E_{data}$  αντιπροσωπεύει το κόστος αντιστοίχισης για τις συντεταγμένες  $(x, y)$ ,  $E_{smooth}$  αντιπροσωπεύει την ενέργεια ομαλότητας και οδηγεί τα γειτονικά *pixel* να έχουν παρόμοιες ανισότητες, οι οποίες βασίζονται στις προηγούμενες παραδοχές και  $\beta$  είναι ένας παράγοντας στάθμισης.

Ο σκοπός της **βελτίωση στερεοσκοπικών διαφορών** είναι να μειώσει τον θόρυβο που υπάρχει στον χάρτη ανισότητας. Συνήθως, το βήμα αυτό έχει να κάνει να με την κανονικοποίηση και την παρεμβολή κάποιων παραμέτρων. Η κανονικοποίηση θα μειώσει τον συνολικό θόρυβο, μέσω ενός φιλτραρίσματος των pixel και με μικρές παραλλαγές μεταξύ των pixel στον χάρτη ανισοτήτων.

### 2.7.2 Προβλήματα στερεοσκοπικής αντιστοιχίας

Η διαδικασία υπολογισμού των αντιστοιχιών μεταξύ των σημείων είναι απαραίτητη προϋπόθεση για την ανάκτηση πληροφοριών σχετικά με το βάθος. Τα εικονοστοιχεία πρέπει να βρεθούν και στις δυο εικόνες και στην συνέχεια να αναλυθούν γεωμετρικά.

Παρόλο που οι δυο εικόνες είναι σχετικά όμοιες, δηλαδή στη μια εικόνα περιέχονται σημεία που είναι όμοια και στην άλλη εικόνα, πολλά προβλήματα προκύπτουν κατά την αναζήτηση και την αντιστοίχιση των pixel στις εικόνες. Γι' αυτό το λόγο υπάρχει μεγάλος αριθμός προτεινόμενων μεθόδων, οι οποίες προσπαθούν να επιλύσουν κάποια από τα προβλήματα που παρουσιάζονται παρακάτω.

- **Φωτομετρική παραμόρφωση(photometric distortion):** Μερικά αντικείμενα λόγω της ανακλαστικής του ιδιότητας, ανακλούν το φως με τον ίδιο τρόπο σε όλη την επιφάνεια τους. Αυτό έχει ως αποτέλεσμα, κάποια σημεία να φαίνονται πιο φωτεινά ανάλογα με από τη θέση που παρατηρούνται από τη κάμερα. Αυτό έχει ως συνέπεια να είναι η δύσκολη η αντιστοίχιση αυτών των σημείων, διότι έχουν σχετικά μεγάλες διαφορές. Στην παρακάτω φωτογραφία φαίνεται το πρόβλημα όπως περιγράφηκε.



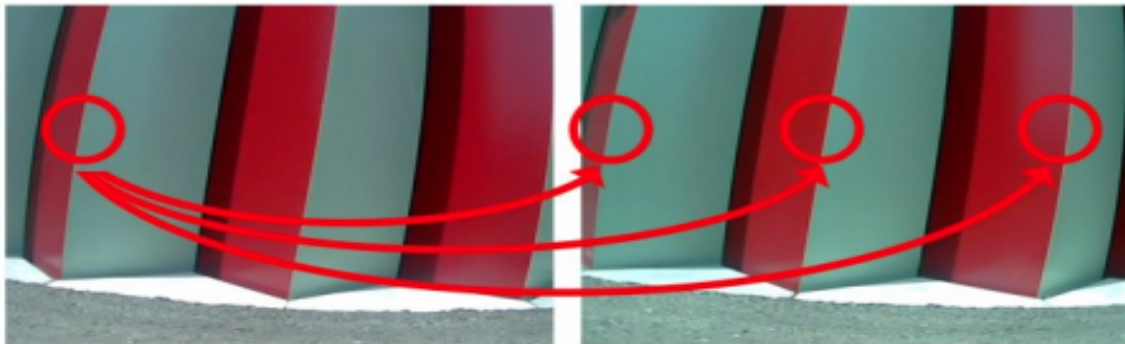
Σχήμα 22: Πρόβλημα στερεοσκοπικής αντιστοίχισης: Φωτομετρική παραμόρφωση [16]

- **Προοπτική παραμόρφωση (projective distortion):** Στην περίπτωση όπου οι εικόνες δεν έχουν διορθωθεί σωστά, κάποια σημεία δεν θα είναι στη ίδια θέση. Με αποτέλεσμα να μην μπορεί να εφαρμοστεί σωστά η αντιστοίχιση.



Σχήμα 23: Πρόβλημα στερεοσκοπικής αντιστοίχησης: Προοπτική παραμόρφωση [16]

- **Ψευδείς αντιστοιχίες (false correspondences):** Συχνό είναι το φαινόμενο, η αντιστοίχιση ενός σημείου της μιας εικόνας να γίνεται με ένα άλλο σημείο, το οποίο δεν αντιπροσωπεύει το τρισδιάστατο σημείο του χώρου. Αυτό συμβαίνει συνήθως σε εικόνες που υπάρχουν περιοδικότητες, όπως στο παρακάτω σχήμα.



Σχήμα 24: Πρόβλημα στερεοσκοπικής αντιστοίχησης: Ψευδείς αντιστοιχίες [16]

- **Αποκλεισμένες περιοχές (occlusions):** Δεν είναι λίγες οι φορές, που παρότι υπάρχει μεγάλη επικάλυψη μεταξύ των εικόνων, υπάρχουν σημεία τα οποία φαίνονται μόνο από την μια από τις δυο κάμερες. Έτσι, είναι σχεδόν αδύνατον να υπάρξει αντιστοιχία. Στην παρακάτω εικόνα μπορεί να γίνει αντιληπτό το φαινόμενο αυτό.

Παραλού τις δυσκολίες που εμφανίζονται κατά την στερεοσκοπική αντιστοίχιση υπάρχουν κάποιοι περιορισμοί που εφαρμόζονται ώστε να ελαχιστοποιηθεί, όσο το δυνατόν γίνεται, η δυσκολία της αντιστοίχιας. Παρακάτω θα αναλυθούν κάποιοι από τους βασικούς περιορισμούς που λαμβάνονται υπόψη.



Σχήμα 25: Πρόβλημα στερεοσκοπικής αντιστοίχισης: Αποκλεισμένες περιοχές [16]

- **Μοναδικότητα:** Υπάρχει μονό μια δυνατή αντιστοιχία μεταξύ των σημείων στις εικόνες.
- **Επιπολικός περιορισμός:** Ο περιορισμός αυτός είναι πολύ σημαντικός, όπως έχει ήδη αναλυθεί στο κεφάλαιο που σχετίζεται με την επιπολική γεωμετρία, διότι η αναζήτηση της αντιστοιχίας των σημείων περιορίζεται μονό στην επιπολική ευθεία. Με αποτέλεσμα να μειώνεται αρκετά το κόστος της υπολογιστής ισχύς και ο χρόνος εκτέλεσης. Ο περιορισμός αυτός για να εφαρμοστεί αναγκαίο κρίνεται να έχει γίνει στερεοσκοπική ανόρθωση στις εικόνες.
- **Περιορισμός συνέχειας:** Ο περιορισμός αυτός αναφέρεται στην ομαλότητα της στερεοσκοπικής διάφορα που προκύπτει. Πιο συγκεκριμένα, δεν πρέπει να υπάρχουν μεγάλες και διακρατικές αλλαγές, άλλα συνεχείς και ομαλές μεταβολές.

Υπάρχουν και άλλοι περιορισμοί που λαμβάνονται κατά τη διαδικασία της στερεοσκοπικής αντιστοίχισης. Ωστόσο, έχουν αναπτυχθεί αρκετοί αλγόριθμοι και τεχνικές για τη στερεοσκοπική αντιστοίχιση. Παρακάτω θα αναλυθούν οι αλγόριθμοι που χρησιμοποιήθηκαν στα πλαίσια της διπλωματικής εργασίας.

### 2.7.3 Κατηγορίες αλγόριθμων στερεοσκοπικής συσχέτισης

Για την αντιμετώπιση του προβλήματος της στερεοσκοπικής συσχέτισης έχουν αναπτυχθεί αρκετοί αλγόριθμοι. Οι αλγόριθμοι αυτοί μπορούν να ταξινομηθούν σε δυο βασικές κατηγορίες ανάλογα με τον τρόπο που βασίζονται ώστε να πετύχουν τη συσχέτιση των σημείων. Η πρώτη κατηγορία είναι οι μέθοδοι που βασίζονται στα χαρακτηριστικά σημεία, ενώ στην άλλη κατηγορία συμπεριλαμβάνονται αλγόριθμοι που είναι βασισμένοι στην συσχέτιση σημείων και περιοχών.

- **Αλγόριθμοι βασισμένοι στα χαρακτηριστικά σημεία (feature-based algorithms):** Οι μέθοδοι αυτοί βασίζονται στην εξαγωγή χαρακτηριστικών σημείων από τις εικόνες και στην συνέχεια προσπαθούν να τα αντιστοιχίσουν μεταξύ τους. Σχετικά είναι μια αποτελεσματική τεχνική, καθώς επιτυγχάνει καλές συσχετίσεις, όμως ο χάρτης βάθους που προκύπτει είναι σχετικά αραιός και θα πρέπει να εφαρμοστεί κάποια μέθοδος για να βελτιωθούν τα αποτελέσματα.
- **Αλγόριθμοι βασισμένοι σε περιοχές (area-based algorithms):** Οι τεχνικές αυτές βασίζονται σε περιοχές-παράθυρα με σταθερό μέγεθος που θα χρησιμοποιηθούν για την αντιστοίχιση. Το κριτήριο ομοιότητας είναι το μέτρο της συσχέτισης ανάμεσα στις περιοχές. Πιο συγκεκριμένα, οι μέθοδοι αυτοί επιλέγουν μια περιοχή γύρω από κάθε εικονοστοιχείο της πρώτης εικόνας και αναζητούν το αντίστοιχο παράθυρο στην δεύτερη εικόνα με την μεγαλύτερη ομοιότητα, σύμφωνα με το μέτρο συσχέτισης. Στην περίπτωση που στις εικόνες έχει εφαρμοστεί στερεοσκοπικής ανόρθωσης, οι μέθοδοι αυτοί πετυχαίνουν πολύ καλά αποτελέσματα με σχετικά μικρό υπολογιστικό κόστος, καθώς η αναζήτηση περιοχών γίνεται πάνω στην επιπολική ευθεία που έχει σχηματιστεί. Ένα από τα βασικά μειονεκτήματα είναι ότι οι αλγόριθμοι αυτοί δεν είναι τόσο αποτελεσματική σε εικόνες που δεν υπάρχει αρκετή υφή, ώστε να δημιουργηθούν ισχυρές συσχετίσεις.

Ένας από τους πιο γνωστούς και αποτελεσματικούς αλγορίθμους που ανήκει στην κατηγορία area-based algorithms είναι ο αλγόριθμος Stereo Block Matching. Αξίζει να αναφερθεί στα πλαίσια της διπλωματικής για τη στερεοσκοπική συσχέτιση χρησιμοποιήθηκαν και συγκρίθηκαν αλγόριθμοι από αυτήν την κατηγορία. Παρακάτω θα αναλυθούν οι αλγόριθμοι που χρησιμοποιήθηκαν στη διπλωματική εργασία.

### 2.7.4 Stereo Block Matching

Ο αλγόριθμος Stereo Block Matching, όπως έχει αναφερθεί και παραπάνω, ανήκει στην κατηγορία area-based algorithms. Η μέθοδος απαρτίζεται από τρία βασικά βήματα.

εγ1: Αρχικά, οι εικόνες φιλτράρονται έτσι ώστε να κανονικοποιηθεί η φωτεινότητα και να αυξηθεί η υφή τους. Πιο συγκεκριμένα, η φωτεινότητα  $I_c$  κάθε εικονο-

στοιχείου αλλάζει σύμφωνα με την παρακάτω σχέση:

$$\min(\max(I_c - \bar{I}, -I_{cap}), I_{cap})$$

οπού,  $\bar{I}$  είναι η μέση τιμή της φωτεινότητας της εκατοστέ γειτονίας, συνήθως το μέγεθος του παραθύρου είναι  $(5 \times 5)$  μέχρι και  $(21 \times 21)$ ,  $I_{cap}$  είναι ένα προκαθορισμένο όριο. **Block Matching Stage:** Σε αυτό βήμα, ο αλγόριθμος αναζητάει τις αντιστοιχίες ολισθαίνοντας ένα μπλοκ στις εικόνες χρησιμοποιώντας ένα κριτήριο ομοιότητας. Υπάρχουν διάφορα κριτήρια ομοιότητας για την σύγκριση ενός παραθύρου με διαστάσεις  $(2a + 1) \times (2b + 1)$ , το οποίο είναι κεντραρισμένο γύρω από τα σημεία  $(u, v)$  της πρώτης εικόνας και  $(u', v')$  της δεύτερης εικόνας.

Τα κάποια από τα κριτήρια ομοιότητας που μπορεί να χρησιμοποιηθούν είναι:

1. • **Sum of Absolute Differences (SAD):** Το κριτήριο αυτό χρησιμοποιείται κατά κύριο λόγο από τον αλγόριθμο. Μαθητικά το κριτήριο αυτό ορίζεται ως:

$$SAD : \sum_{k=-a}^b \sum_{l=-b}^b [I_1(u + k, v + 1) - I_2(u' + k, v' + 1)]$$

- **Sum of Squared Differences:**

$$\sum_{k=-a}^b \sum_{l=-a}^b [I_1(u + k, v + 1) - I_2(u' + k, v' + 1)]^2$$

Υπάρχουν και άλλα κριτήρια ομοιότητας που μπορούν να χρησιμοποιηθούν όπως τα κριτήρια Normalized Cross Correlation, Rank, Census.

2. **Postfiltering stage:** Στο τελικό στάδιο εφαρμόζεται ένα φίλτρο στις υποψήφιες αντιστοιχίες που έχουν βρεθεί στο προηγούμενο στάδιο, κατά το οποίο οι μη σαφές ορισμένες αντιστοιχίες αποκλείονται. Πιο συγκεκριμένα, σε περίπτωση που η μέγιστη τιμή του μέτρου ομοιότητας δεν είναι αρκετά μεγαλύτερη της ελάχιστης τιμής τότε δεν λαμβάνονται υπόψη.

### 2.7.5 Semi Global Block Matching

Ο αλγόριθμος αυτός αναπτύχθηκε από τον Hirschmuller και βασίζεται στη συσχέτιση περιοχών-παραθύρων. Ο Semi Global Block Matching αποτελείται από τα εξής παρακάτω βήματα.

1. **Υπολογισμός του κόστους αντιστοιχίας (Cost computation):** Σε αυτό το αρχικό στάδιο, επιλέγεται μια μέγιστη τιμή για την στερεοσκοπική διάφορα  $d_{max}$  και σαν είσοδο δίνονται δυο εικόνες στις οποίες έχει εφαρμοστεί στερεοσκοπική ανόρθωση. Ως εικόνα βάσης ορίζεται η αριστερή εικόνα  $I_l$ , για κάθε  $pixel(x, y)$  της εικόνας βάσης είναι βέβαιο ότι το αντίστοιχο  $pixel$  της άλλης εικόνας θα βρίσκεται μεταξύ των θέσεων  $(x, y)$  και  $(x - d_{max}, y)$ . Οι τιμές για τη στερεοσκοπική διάφορα αξιολογούνται με βάση τη συνάρτηση κόστους  $C_{BT}(p, d)$ :

$$C_{BT}(p, d) = [I_l(p) - I_r(p - d)]$$

2. **Άθροισμα του κόστους (Cost aggregation):** Για να αποφευχθούν οι λανθασμένες και να βρεθούν οι σωστές στερεοσκοπικές διαφορές που προκύπτουν από την συνάρτηση κόστους, χρησιμοποιείται η ενέργεια της εικόνας στερεοσκοπικών διαφορών. Πιο συγκεκριμένα, η ενέργεια των στερεοσκοπικών διαφορών υπολογίζεται από το σύνολο όλων των κοστών από των εικονοστοιχείων για τη στερεοσκοπική διαφορά  $d$ . Επίσης, ο αλγόριθμος χρησιμοποιεί τα αρχικά κόστη και άλλες δυο παραμέτρους, οι οποίες εξαρτώνται από τη διάφορα των αποκλίσεων σε μια περιοχή. Στην περίπτωση όπου οι αποκλίσεις διαφέρουν κατά 1 pixel προστίθεται η παράμετρος  $P_1$  με τα κόστη. Η παράμετρος αυτή εξασφαλίζει την ομαλότητα των περιοχών. Για την περίπτωση όπου οι διαφορές είναι μεγαλύτερες προστίθεται η παράμετρος  $P_2$ . Μαθηματικά ισχύει ότι:

$$E(D) = \sum_P C_{BT}(p, d_p) + \sum_q P_1 T[|D_p - D_q| = 1] + \sum_q P_2 T[|D_p - D_q| > 1]$$

οπού,

το  $p$  ανήκει στην περιοχή που επιλέγεται. Ο αλγόριθμος στοχεύει στην ελαχιστοποίηση της συνάρτησης ενέργειας  $E(D)$ . Η ελαχιστοποίηση της συνάρτησης αυτής είναι ένα σχετικά δύσκολο πρόβλημα, καθώς είναι ένα διδιάστατο πρόβλημα. Όμως ο αλγόριθμος για να απλοποιήσει το πρόβλημα αυτό, διασχίζονται μονοδιάστατα μονοπάτια  $L_r$ , στα οποία εφαρμόζονται και κάποιοι περιορισμοί. Για κάθε διαδρομή, το ελάχιστο κόστος για να φτάσει στο εικονοστοιχείο  $p$  υπολογίζεται από τη συνάρτηση  $E(D)$ . Η συνάρτηση  $L_r$  είναι αναδρομική και ορίζει άλλο ένα μέτρο κόστους για κάθε  $pixel$  σε μια συγκεκριμένη θέση. Επομένως, για να υπολογιστεί το κόστος για ένα εικονοστοιχείο χρειάζονται όλες τις τιμές κόστους

των προηγούμενων pixels. Η διάσχιση του μονοπατιού γίνεται από αριστερά προς δεξιά. Είναι δυνατόν ένα εικονοστοιχείο να έχει χαμηλότερη τιμή κόστους εάν ληφθεί υπόψη μονό η γειτονία από μια διαδρομή. Κρίνεται αναγκαίο να επαναληφθεί η διαδικασία για διαφορετικά μονοπάτια με διαφορετικούς προσανατολισμούς. Ο ελάχιστος αριθμός των διαδρομών είναι οκτώ.

$$L_r(p, d) = C_{BT}(p, D_p) + \min(L_r(p - r, d), L_r(p - r, d - 1) + P_1, L_r(p - r, d + 1)) + P_1, \min_i L_r(p - r, i) + P_2) - \min_k L_r(p - r, k)$$

Στη γενική μορφή η παραπάνω συνάρτηση μπορεί να γραφτεί

$$S(p, D) = \sum_r L_r(p, d)$$

Η σχέση αυτή μας παρέχει τη συνολική πληροφορία από όλα τα μονοπάτια.

### 3. Υπολογισμός στερεοσκοπικών αποκλίσεων (Disparity computation)

Σε αυτό το τελευταίο βήμα υπολογίζεται η στερεοσκοπική διάφορα, η οποία για κάθε εικονοστοιχείο ελαχιστοποιεί το κόστος, χρησιμοποιώντας ως δεδομένο τη συνάρτηση  $S$ . Πιο συγκεκριμένα, η τιμή  $D_L$ , για κάθε εικονοστοιχείο είναι:

$$D_L(p) = \operatorname{argmin}_d (S(p, d))$$

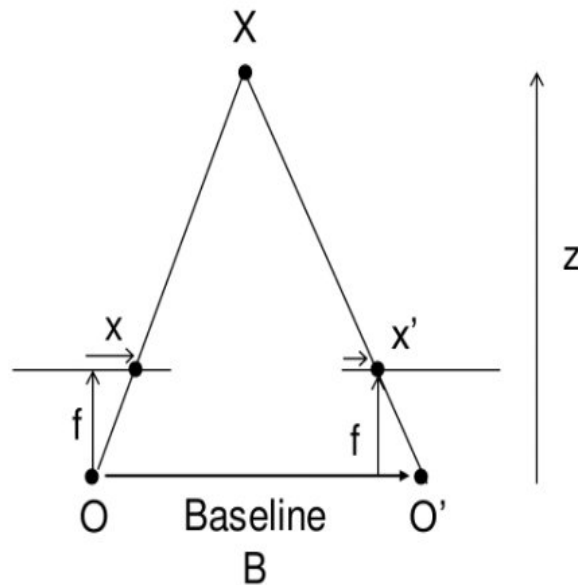
Ο χάρτης διάφορων  $D_L$  διαμορφώνεται επιλέγοντας τη διαδρομή με το ελάχιστο κόστος για κάθε εικονοστοιχείο. Αυτό συμβαίνει μέσω της subpixel accuracy επιλέγοντας το ελάχιστο μιας τετραγωνικής καμπύλης, μέσω του κόστους των γειτονικών εικονοστοιχείων. [5]

## 2.8 Ανακατασκευή σκηνής - Τριγωνοποίηση

Η ανακατασκευή της σκηνής βασίζεται στην **τριγωνοποίηση (triangulation)**. Ουσιαστικά, με τη βοήθεια της τριγωνοποίησης θα υπολογιστεί το βάθος ενός σημείου που απεικονίζεται στο στερεοσκοπικό ζεύγος εικόνων. Σημαντικό είναι να αναφερθεί ότι το στερεοσκοπικό σύστημα πρέπει να έχουν ολοκληρωθεί με επιτυχία όλα τα απαραίτητα βήματα. Όπως η βαθμονόμηση της στερεοσκοπικής κάμερας, η ανόρθωση των στερεοσκοπικών εικόνων, καθώς επίσης να έχει επιτευχθεί στερεοσκοπική συσχέτιση στις εικόνες.

Παίρνοντας ως δεδομένο ότι η στερεοσκοπική κάμερα είναι ιδανική, δηλαδή οι κάμερες απέχουν μεταξύ τους απόσταση  $b$  (baseline), ενώ είναι παράλληλες και ευθυγραμμισμένες με τους οπτικούς άξονες. Επίσης, όλοι οι παράμετροι των καμερών είναι γνωστές και οι εστιακές τους αποστάσεις είναι ίσες.

Υποθέτοντας ένα σημείο  $X(x, y, z)$  της σκηνής και  $x$  το σημείο προβολής του σημείου  $X$  για τη δεξιά κάμερα, ενώ αντίστοιχα  $x'$  για την αριστερή κάμερα. Τα σημεία αυτά έχουν βρεθεί μετά από την εφαρμογή της στερεοσκοπικής αντιστοίχισης. Στο Σχήμα 26 φαίνεται γεωμετρικά το στερεοσκοπικό σύστημα.



Σχήμα 26: Εφαρμογή τριγωνοποίησης για τον υπολογισμό του βάθους [13]

Εφαρμόζοντας όμοια τρίγωνα, προκύπτει η παρακάτω σχέση

$$disparity = x - x' = \frac{Bf}{Z}$$

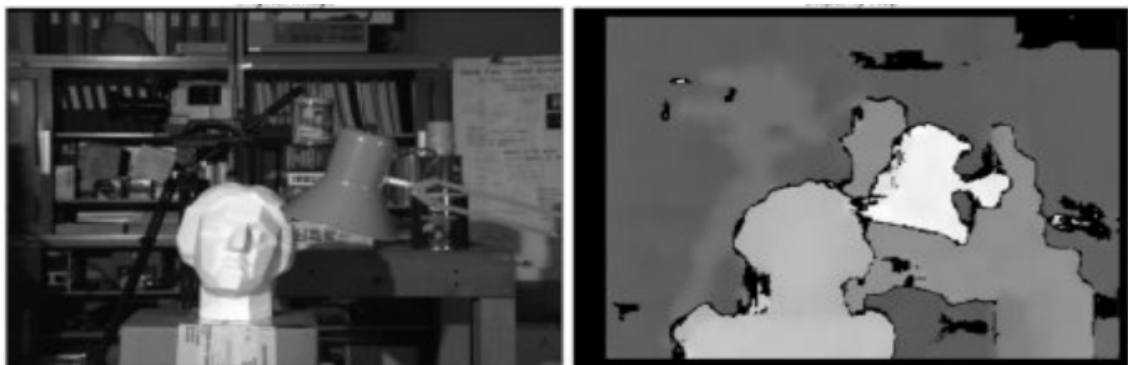
οπού,

- $B$  είναι η απόσταση μεταξύ των καμερών,
- $f$  είναι το εστιακό βάθος των καμερών οπού είναι ίδιο και στις δυο κάμερες
- $x$  και  $x'$  είναι η απόσταση μεταξύ των σημείων της εικόνας απο τα οπτικά κέντρα των καμερών
- $Z$  το βάθος.

Με βάση την παραπάνω σχέση ως *disparity* ορίζεται η διάφορα των δυο προβολών σε *pixels*. Λύνοντας όπως προς  $Z$  τη σχέση αυτή υπολογίζεται εύκολα το βάθος, ισχύει ότι

$$Z = \frac{Bf}{disparity}$$

Με την εφαρμογή της τριγωνοποίησης για όλα τα εικονοστοιχεία της σκηνής προκύπτει ο χάρτης ανισοτήτων (*disparity map*). Πιο συγκεκριμένα, δημιουργείται μια νέα εικόνα στην οποία η έντασή της φωτεινότητας κάθε *pixel* έχει άμεση σχέση με το *disparity*. Όταν τα αντικείμενα είναι κοντά στο στερεοσκοπικό σύστημα φαίνονται πιο έντονα, αντίθετα τα αντικείμενα που είναι πιο μακριά φαίνονται λιγότερα έντονα και πιο σκοτεινά. Στην παρακάτω εικόνα, γίνεται αντιληπτό ότι το φωτιστικό είναι πολύ κοντά στην κάμερα, καθώς είναι το πιο φωτεινό αντικείμενο, αντίθετα με τα άλλα αντικείμενα που βρίσκονται πιο πίσω και έχουν διαφορετικές και πιο σκούρες αποχρώσεις.



Σχήμα 27: Disparity map [13]

Με την εύρεση του χάρτη ανισοτήτων, μπορεί να υπολογιστεί ο χάρτης βάθους (depth map). Έτσι, είναι εφικτό να πραγματοποιηθεί η τρισδιάστατη ανακατασκευή της σκηνής, αφού έχουν υπολογιστεί και οι τρεις συνταγμένες κάθε σημείου. Οπότε οι  $3D$  συντεταγμένες ενός σημείου  $K(x, y, z)$  είναι :

$$K(x, y, z) = (x, y, \frac{bf}{disparity})$$

### 3 Υλοποίηση εφαρμογής

Ένας από τους βασικούς στόχους της διπλωματικής εργασίας είναι η ανάπτυξη και η υλοποίηση μιας πλήρους λειτουργικής εφαρμογής, η οποία θα στοχεύει στην τρισδιάστατη ανακατασκευή εικόνων. Επιπλέον, η εφαρμογή προσφέρει γραφικό περιβάλλον, όπου ο χρήστης έχει την δυνατότητα να συνδυάζει όλες τις απαραίτητες διαδικασίες και λειτουργίες για την επίτευξη του στόχου.

#### 3.1 Βιβλιοθήκη OpenCV

Η OpenCV (Open Source Computer Vision Library) είναι μια ανοιχτού κώδικα βιβλιοθήκη για υπολογιστική όραση και μηχανική μάθηση. Δημιουργήθηκε για να παρέχει μια κοινή υποδομή για εφαρμογές υπολογιστικής όρασης και για να επιταχύνει τη χρήση της μηχανικής αντίληψης στα εμπορικά προϊόντα. Έχει BSD (Berkeley Software Distribution) άδεια καθιστώντας εύκολο για τις επιχειρήσεις να χρησιμοποιούν και να τροποποιούν τον κώδικα. Η OpenCV έχει ενσωματωμένους στις βιβλιοθήκες της, περισσότερους από 2500 βελτιστοποιημένους αλγόριθμους. Περιλαμβάνει, ένα ολοκληρωμένο σύνολο κλασικών αλλά και state-of-the art αλγορίθμων υπολογιστικής όρασης. Αναφορικά κάποιοι από τους αλγορίθμους αυτούς χρησιμοποιούνται σε εφαρμογές που αφορούν:

- την ανίχνευση και την αναγνώριση προσώπων
- τον εντοπισμό στάσιμων και κινούμενων αντικειμένων,
- **την παραγωγή 3D point clouds από στερεοσκοπικά ζεύγη καμερών**
- την εύρεση παρόμοιων εικόνων από μια βάση δεδομένων με εικόνες,
- τον εντοπισμό της κίνησης των ματιών
- τη διεπαφή χρήστη υπολογιστή

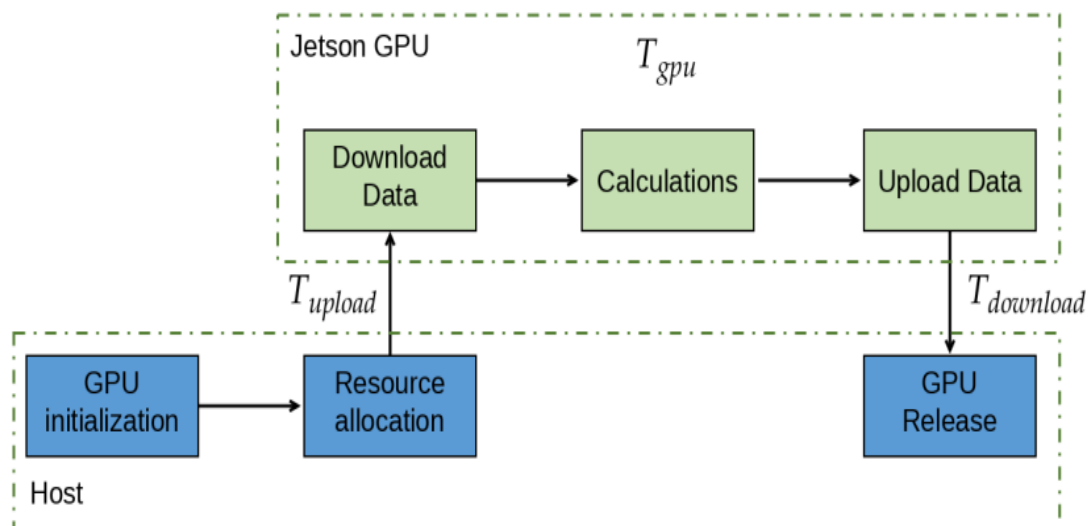
Μεταξύ των διαθέσιμων αλγορίθμων, χρησιμοποιήθηκε το calib3d module, το οποίο περιλαμβάνει βασικούς αλγόριθμους γεωμετρίας πολλαπλής προβολής, βαθμονόμηση απλής και στερεοσκοπικής κάμερας, εκτίμηση πόζας αντικειμένου και αλγόριθμους που για την τρισδιάστατη ανακατασκευή αντικειμένων. [8]

### 3.2 Compute Unified Device Architecture (CUDA)

CUDA είναι μια πλατφόρμα που χρησιμοποιεί παραλληλισμό για τους υπολογισμούς, έχει αναπτυχθεί από την εταιρία NVIDIA για να αξιοποιεί τα δικά της GPUs. Μέσω της πλατφόρμας αυτής, οι προγραμματιστές μπορούν να την ενσωματώσουν στις εφαρμογές που αναπτύσσουν για να αυξήσουν την υπολογιστή ισχύ.

Από το 2010 αναπτύχθηκε το CUDA module για την βιβλιοθήκη OpenCV με σκοπό τη βελτίωση των αλγορίθμων αξιοποιώντας τα NVIDIA GPUs. Σκοπός είναι αλγόριθμοι να μπορούν να χρησιμοποιηθούν σε πραγματικό χρόνο.

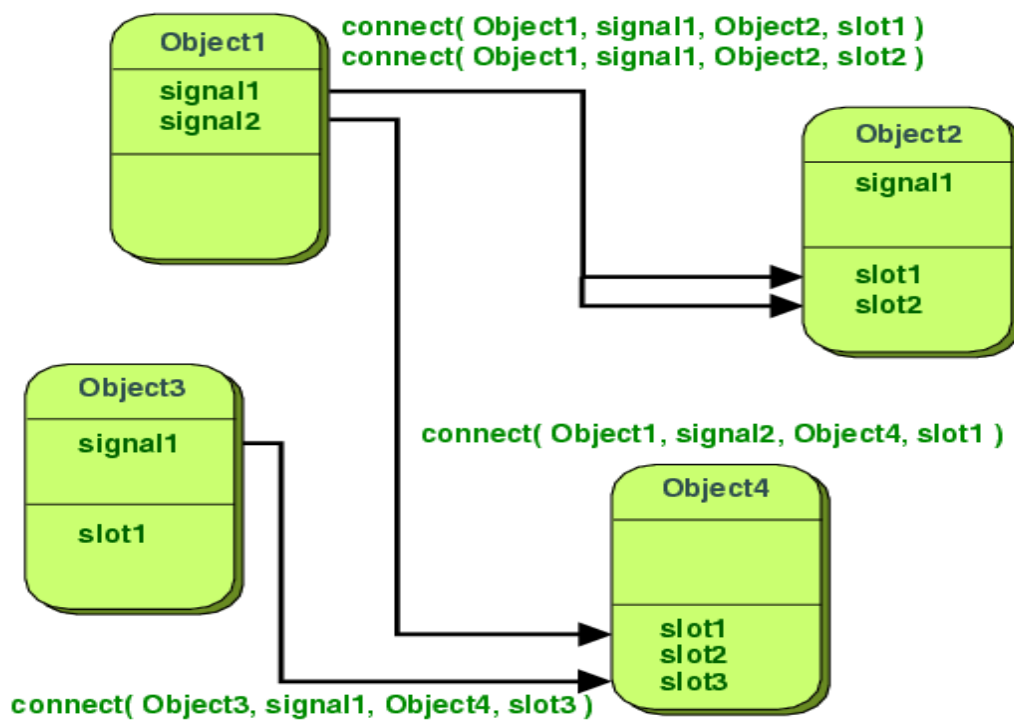
Ο τρόπος επικοινωνίας είναι περίπλοκος. Τα δεδομένα πρέπει να μεταφέρουν από το CPU στα GPU. Αρχικά, γίνεται μια αρχικοποίηση και μια δέσμευση πόρων, στην συνέχεια στέλνονται τα δεδομένα στα GPUs και ξεκινάει η διαδικασία υπολογισμού. Τέλος, τα αποτελέσματα στέλνονται πίσω στο CPU και γίνεται αποδέσμευση των πόρων. Στο παρακάτω σχήμα φαίνεται πως γίνεται η μεταφορά των δεδομένων και ο υπολογισμός τους.



Σχήμα 28: Ποή δεδομένων CPU to GPU

### 3.3 Βιβλιοθήκη Qt

Η Qt είναι μια βιβλιοθήκη και πλατφόρμα για την ανάπτυξη γραφικών εφαρμογών. Η βασική ιδέα της βιβλιοθήκης είναι τα signals and slots. Πιο συγκεκριμένα, τα αντικείμενα που δημιουργούνται από τις κλάσεις Qt μπορούν να επικοινωνούν απευθείας μεταξύ τους. Επίσης, η βιβλιοθήκη έχει υλοποιηθεί σε γλώσσα C++. Είναι σημαντικό να αναφερθεί ότι η βιβλιοθήκη αυτή μπορεί να χρησιμοποιηθεί για την ανάπτυξη εφαρμογής για οποιοδήποτε λειτουργικό σύστημα.



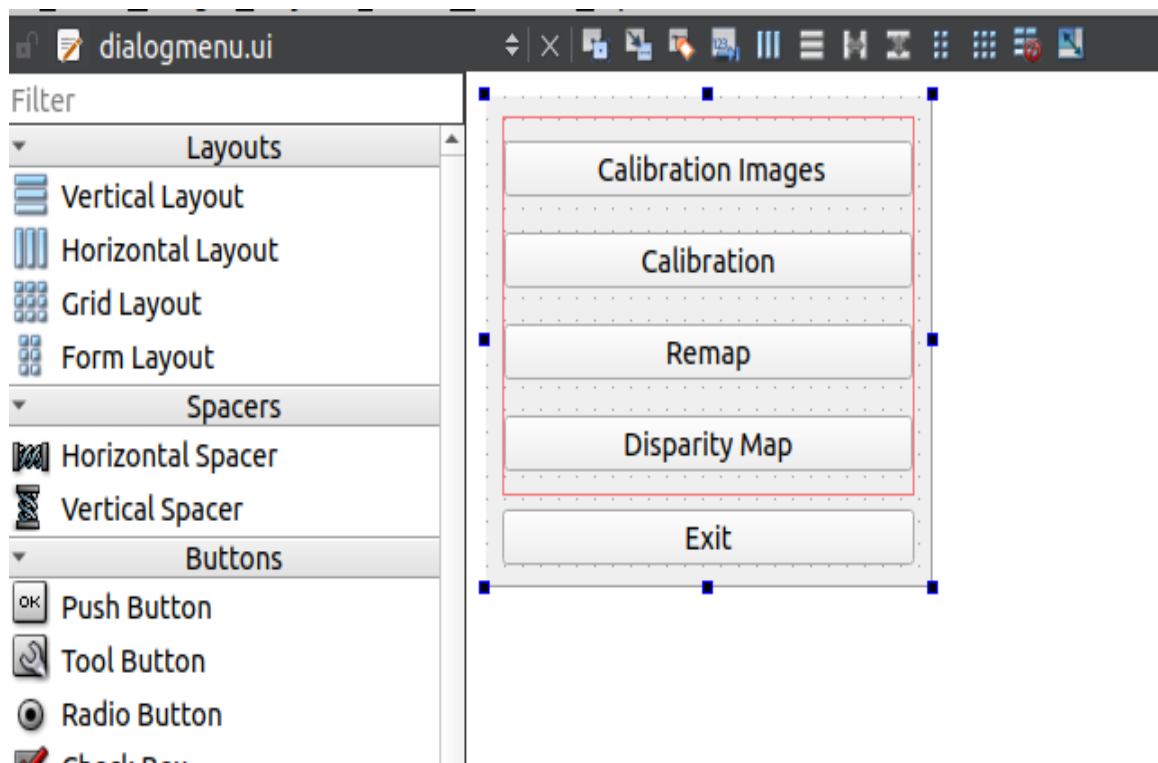
Σχήμα 29: Qt: Singals and Slots

### 3.4 Qt Creator

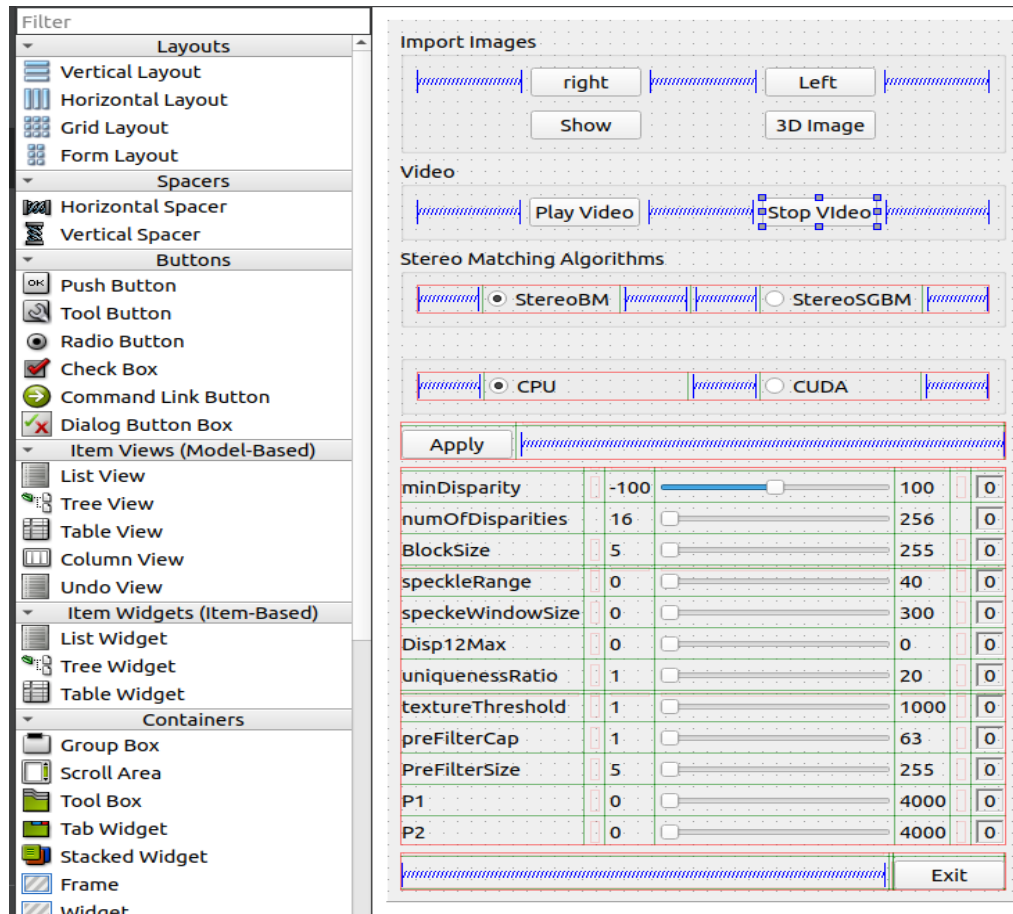
Το Qt Creator είναι ένα ολοκληρωμένο περιβάλλον ανάπτυξης  $C++$ , που απλοποιεί την ανάπτυξη εφαρμογών GUI. Είναι ένα εργαλείο το οποίο χρησιμοποιεί κατά κύριο λόγο τη βιβλιοθήκη *Qt*. Το Qt Creator χρησιμοποιήθηκε ως βασικό εργαλείο για την ανάπτυξη και υλοποίηση της εφαρμογής.

Μέσω του προγράμματος δίνεται η δυνατότητα να αναπτυχθεί εύκολα ένα γραφικό περιβάλλον για την εφαρμογή, καθώς το Qt Creator προσφέρει έτοιμα αντικείμενα για επεξεργασία.

Παρακάτω παρουσιάζονται μερικές εικόνες από τη σχεδίαση της εφαρμογής, μέσω του Qt Creator.



Σχήμα 30: Qt Creator: Menu GUI



Σχήμα 31: Qt Creator: Disparity Map GUI

### 3.5 Υλοποίηση εφαρμογής βαθμονόμησης

Για τη βαθμονόμηση της στερεοσκοπικής κάμερας αναπτύχθηκε μια εφαρμογή, που αποτελεί βασικό στόχο της παρούσας διπλωματικής εργασίας, η οποία αξιοποιεί την βιβλιοθήκη OpenCV. Η διαδικασία βαθμονόμησης που χρησιμοποιείται βασίζεται στην τεχνική βαθμονόμησης του Zhengyou Zhang, δηλαδή στην παρατήρηση ενός δισδιάστατου αντικειμένου από διαφορετικές γωνίες. Πιο συγκεκριμένα, με αυτό τον τρόπο είναι δυνατό να υπολογιστεί η σχετική θέση και ο προσανατολισμός της κάμερας τη στιγμή κάθε λήψης καθώς επίσης και οι εσωτερικοί παράμετροι της κάμερας. Η παροχή πολλαπλών λήψεων επιτυγχάνεται με περιστροφή και μετακίνηση του αντικειμένου. Το διασδιάστατο αντικείμενο που θα χρησιμοποιηθεί είναι μια σκακιέρα.

Δεδομένων των πολλών κοινών λήψεων της σκακιέρας, η διαδικασία βαθμονόμηση γίνεται ξεχωριστά για την κάθε κάμερα. Στη συνέχεια, συνδυάζοντας τις παραμέτρους περιστροφής και μεταφοράς των καμερών υπολογίζεται η σχέση της περιστροφής και της μεταφοράς ανάμεσα στις δυο κάμερες. Πιο συγκεκριμένα, αυτό που συμβαίνει είναι η μήτρα περιστροφής θα τοποθετήσει την δεξιά κάμερα στο ίδιο επίπεδο με την αριστερή κάμερα.

Η εφαρμογή δίνει τη δυνατότητα στον χρήστη να τραβήξει φωτογραφίες από τη στερεοσκοπική κάμερα και στη συνέχεια αυτόματα να υπολογιστούν οι παράμετροι της βαθμονόμησης. Η διαδικασία που πρέπει να ακολουθηθεί έχει ήδη περιγράψει αναλυτικά στα προηγούμενα κεφάλαια. Παρακάτω θα αναλυθούν τα βασικά βήματα που πρέπει να ακολουθήσει ο χρήστης για να βαθμονομήσει τη στερεοσκοπική κάμερα μέσω της εφαρμογής.

Αρχικά πρέπει να δημιουργηθεί ένα δισδιάστατο αντικείμενο, τύπου σκακιέρας, του οποίου οι διαστάσεις πρέπει να είναι γνωστές και με ακρίβεια ορισμένες. Είναι πολύ εύκολο να κατασκευαστεί το αντικείμενο αυτό, καθώς το μόνο που απαιτείται είναι η αποτύπωσή του πάνω σε μια επιφάνεια, όπως χαρτί. Επίσης, μπορεί να δημιουργηθεί μέσω ενός σχεδιαστικού προγράμματος και στη συνέχεια να εκτυπωθεί. Στο Σχήμα 15 απεικονίζεται το μοτίβο του αντικειμένου.

Στη συνέχεια ο χρήστης πρέπει να ακολουθήσει τα εξής βασικά βήματα.

### Λήψη φωτογραφιών αντικειμένου βαθμονόμησης

1. Ο χρήστης πρέπει να τρέξει την εντολή 'v4l2-ctl --list-devices' στο terminal, ώστε να βρει τα path για τις κάμερες που θέλει να βαθμονομήσει.

```
benmalef@benmalef-desktop:~$ v4l2-ctl --list-devices
Video Capture 3 (usb-0000:00:14.0-5):
    /dev/video0
    /dev/video1

Video Capture 3 (usb-0000:00:14.0-6):
    /dev/video2
    /dev/video3
```

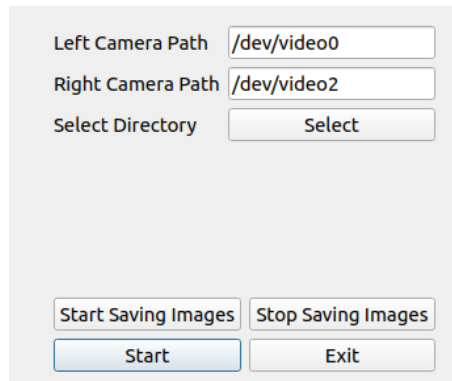
Σχήμα 32: Εντολή στο terminal για την τοποθεσία των καμερών στο υπολογιστή

2. Ο χρήστης για τραβήξει φωτογραφίες του αντικείμενου βαθμονόμησης, πρέπει να επιλέξει από το μενού της εφαρμογής την επιλογή Calibration Images.

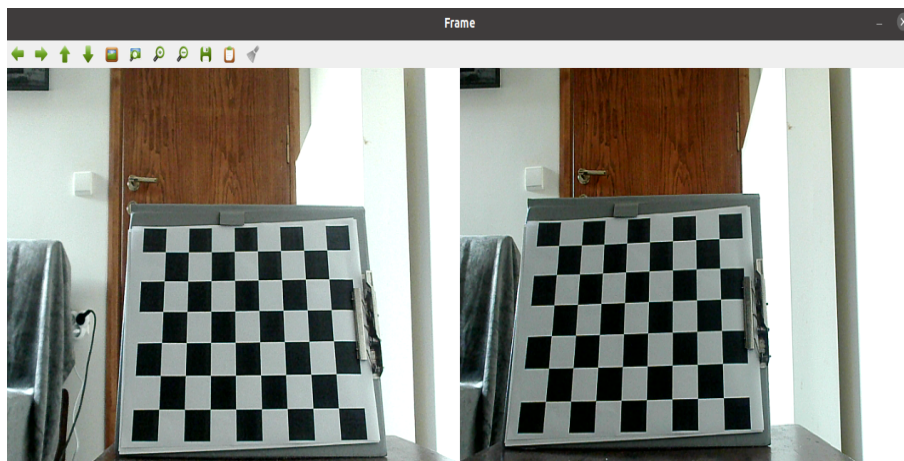


Σχήμα 33: Εφαρμογή για την βαθμονόμηση της στερεοσκοπικής κάμερας.

3. Ο χρήστης πρέπει να εισάγει στα κατάλληλα πεδία τις τοποθεσίες των καμερών στον υπολογιστή. Έπειτα πρέπει να επιλέξει που θέλει να αποθηκεύσει τις φωτογραφίες στον υπολογιστή, πατώντας το κουμπί 'Select'.
4. Πατώντας το κουμπί 'Start' θα ανοίξουν και οι δυο κάμερες.



Σχήμα 34: Εφαρμογή για φωτογράφιση του αντικειμένου βαθμονόμησης.

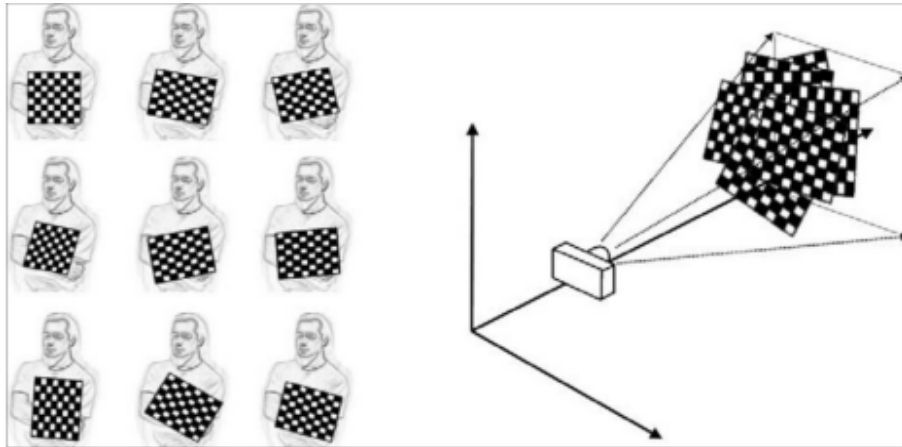


Σχήμα 35: Εφαρμογή για φωτογράφιση του αντικειμένου βαθμονόμησης.

5. Πατώντας το κουμπί 'Start Saving Images' θα αρχίσουν να αποθηκεύονται φωτογραφίες ανά 3 δευτερόλεπτα. Ο χρόνος έχει επιλεγεί έτσι ώστε να προλαβαίνει ο χρήστης να ετοιμάζει το αντικείμενο βαθμονόμησης για την επομένη λήψη του. Πατώντας το κουμπί 'Stop Saving Images' τερματίζεται η διαδικασία λήψης των φωτογραφιών.

Ο χρήστης χρειάζεται να περιστρέφει και να κινεί το αντικείμενο βαθμονόμησης ώστε να καταφέρει πολλαπλές λήψεις του για καλύτερα αποτελέσματα. Στο Σχήμα 36 γίνεται εύκολα αντιληπτό πως πρέπει να γίνουν οι λήψεις.

6. Πατώντας το κουμπί 'Exit' ο χρήστης επιστρέφει στο αρχικό μενού.

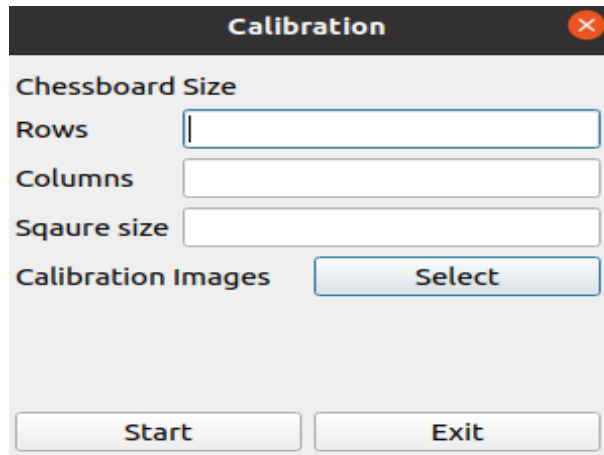


Σχήμα 36: Ενδεικτικές φωτογραφίες της σκακιέρας. [8]

### Βαθμονόμηση στερεοσκοπικής κάμερας

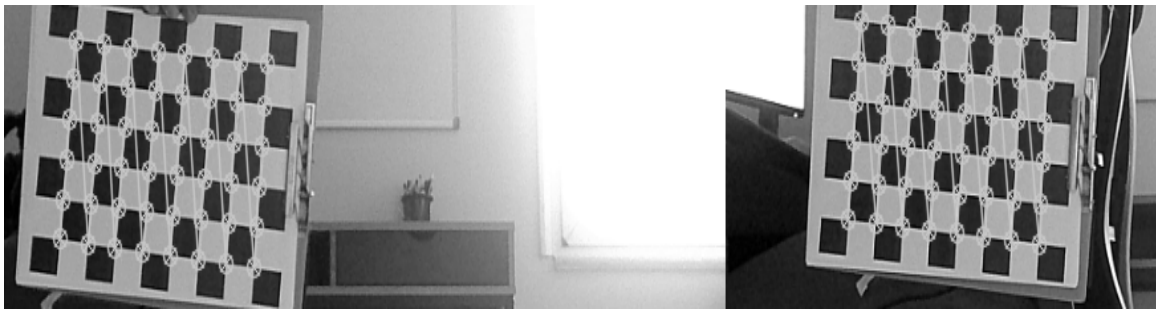
1. Ο χρήστης πρέπει να επιλέξει την επιλογή Calibration. Στη συνέχεια θα του εμφανιστεί μια φόρμα που αφορά κάποιες παραμέτρους σχετικά με την σκακιέρα.
2. Ο χρήστης πρέπει να συμπληρώσει τα πεδία της φόρμας που άφορα τις παραμέτρους της σκακιέρας. Το πεδίο Rows αναφέρεται στις γραμμές από τις οποίες αποτελείται η σκακιέρα. Το πεδίο Columns αναφέρεται στις στήλες της σκακιέρας. Το πεδίο Square Size αναφέρεται στο μέγεθος των τετραγώνων που αποτελούν τη σκακιέρα, μετρημένο σε χιλιοστά.
3. Ο χρήστης πρέπει να εισάγει την τοποθεσία που είναι αποθηκευμένες οι εικόνες του αντικειμένου βαθμονόμησης.
4. Πατώντας το κουμπί 'Start' θα ξεκινήσει η διαδικασία βαθμονόμησης.

Κατά τη διαδικασία της βαθμονόμησης εντοπίζονται τα κατάλληλα σημεία στα ζευγάρια φωτογραφιών του αντικειμένου παρατήρησης. Αν και στις δυο φωτογραφίες εντοπιστούν τα σημεία τότε χρησιμοποιούνται για τον υπολογισμό των παραμέτρων της βαθμονόμησης, αλλιώς προχωράμε στο επόμενο ζευγάρι. Στην παρακάτω φωτογραφία απεικονίζεται πως πρέπει να είναι ζεύγος φωτογραφιών, ώστε να ληφθεί υπόψη στη διαδικασία βαθμονόμησης. Κατά τη διαδικασία της βαθμονόμησης, υπολογίζεται και το σφάλμα βαθμονόμησης τόσο για την κάθε κάμερα ξεχωριστά, όσο και για τη στερεοσκοπική κάμερα. Ανάλογα με το σφάλμα πρέπει να επαναληφθεί η διαδικασία, αν το σφάλμα είναι κοντά στο μηδέν δεν χρειάζεται, σε κάθε άλλη περίπτωση κάλλο είναι να επαναληφθεί.



Σχήμα 37: Ενδεικτικές φωτογραφίες της σκακιέρας

Μετά το τέλος της βαθμονόμησης, όλες οι παράμετροι έχουν αποθηκευτεί σε ένα αρχείο, το οποίο μπορεί να χρησιμοποιηθεί για οποιαδήποτε άλλη εφαρμογή απαιτεί βαθμονόμηση της στερεοσκοπικής κάμερας.



Σχήμα 38: Εφαρμογή για φωτογράφιση του αντικειμένου βαθμονόμησης.

5. Πατώντας το κουμπί 'Exit' ο χρήστης επιστρέφει το αρχικό μενού.

### 3.6 Μέθοδος ανόρθωσης (Rectification method)

Για την υλοποίηση της στερεοσκοπικής ανόρθωσης χρησιμοποιήθηκαν δυο πολύ σημαντικές συναρτήσεις από τη βιβλιοθήκη ανοικτού κωδικά OpenCV. Η πρώτη συνάρτηση ονομάζεται **initUndistortRectifyMap**, ενώ η δεύτερη ονομάζεται **remap**. Η υλοποίηση γίνεται χωρίζεται σε δυο μέρη για να απλοποιηθεί η διαδικασία της ανόρθωσης εικόνων. Θεωρώντας ότι δεν αλλάζουν οι παράμετροι των καμερών και οι κάμερες είναι ίδιες και παραμένουν στον ίδιο προσανατολισμό και θέση. Η πρώτη συνάρτησή υπολογίζει τις παραμέτρους ανόρθωσης για την αριστερή και δεξιά των εικόνων για τις αντίστοιχες κάμερες, η δεύτερη συνάρτηση εφαρμόζει τις υπολογισμένες παραμέτρους στις εικόνες. Ουσιαστικά, είναι το τελικό στάδιο για την εφαρμογή της στερεοσκοπικής ανόρθωσης. Παρακάτω θα αναλυθούν οι συναρτήσεις τόσο σε μαθηματικό επίπεδο, όσο και σε επίπεδο υλοποίησης.

#### 3.6.1 InitUndistortRectifyMap

Η συνάρτηση InitUndistortRectifyMap υπολογίζει τις παραμέτρους για την διόρθωση των παραμορφώσεων που έχουν προκληθεί και τις παραμέτρους για την ανόρθωση των εικόνων. Ουσιαστικά, η συνάρτηση επιστρέφει παραμέτρους που αφορούν χάρτες ανόρθωσης των εικόνων. Αυτοί οι χάρτες δείχνουν που απαιτείται παρεμβολή των pixel της εικόνας. Η συνάρτηση καλείται ξεχωριστά για την κάθε κάμερα. Οι παράμετροι που χρειάζεται η συνάρτηση για να υπολογίσει τους χάρτες είναι οι εσωτερικές παράμετροι της κάμερας, οι παράμετροι για τις παραμορφώσεις, ο πίνακας περιστροφής και ο πίνακας προβολής. Για να δημιουργηθούν οι χάρτες αυτοί, οι οποίοι έπειτα θα χρησιμοποιηθούν από την συνάρτηση remap, για κάθε εικονοστοιχείο  $(u, v)$  στη διορθωμένη εικόνα υπολογίζει τις αντίστοιχες συντεταγμένες στην αρχική εικόνα. Πιο συγκεκριμένα, εφαρμόζεται η παρακάτω διαδικασία: [8]

$$\begin{aligned}x &\leftarrow \frac{u - c_x}{f'_x} \\y &\leftarrow \frac{v - c_y}{f'_y} \\[XYW]^T &\leftarrow R^{-1} \times [xy1]^T \\x' &\leftarrow \frac{X}{W} \\y' &\leftarrow \frac{Y}{W} \\x'' &\leftarrow x'(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + 2p_1 x' y' + p_2(r^2 + 2x'^2) \\y'' &\leftarrow y'(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + p_1(r^2 + 2y'^2 + 2p_2 x' y')\end{aligned}$$

$$\begin{aligned} \text{map}_x(u, v) &\leftarrow x''f_x + c_x \\ \text{map}_y(u, v) &\leftarrow y''f_y + c_y \end{aligned}$$

### 3.6.2 Remap

Η συνάρτηση `remap` εφαρμόζει έναν γεωμετρικό μετασχηματισμό σε μια εικόνα, δηλαδή μετασχηματίζει την αρχική εικόνα σύμφωνα με τους συγκεκριμένους χάρτες που έχουν υπολογιστεί από την παραπάνω συνάρτηση. Ο μετασχηματισμός που εφαρμόζεται είναι:

$$\text{dst}(x, y) = \text{src}(\text{map}_x(x, y), \text{map}_y(x, y))$$

Η συνάρτηση αυτή παίρνει σαν ορίσματα:

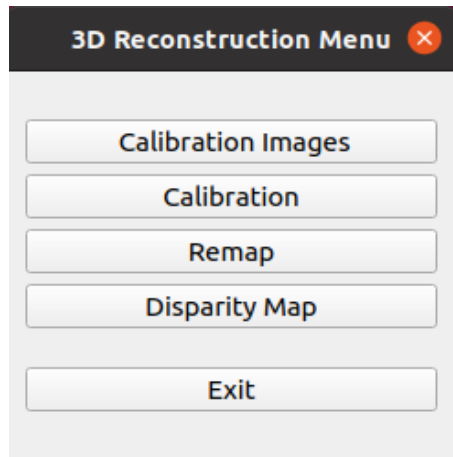
- την αρχική εικόνα
- τους χάρτες που έχουν υπολογιστεί από την συνάρτηση `InitUndistortRectifyMap`
- την εικόνα που θα εφαρμοστεί η διαδικασία ανόρθωσης.

Πρέπει να αναφερθεί ότι η εικόνα στην οποία θα εφαρμοστεί η ανόρθωση πρέπει να έχει ίδιο μέγεθος με την αρχική εικόνα. [8]

### 3.6.3 Υλοποίηση στερεοσκοπικής ανόρθωσης

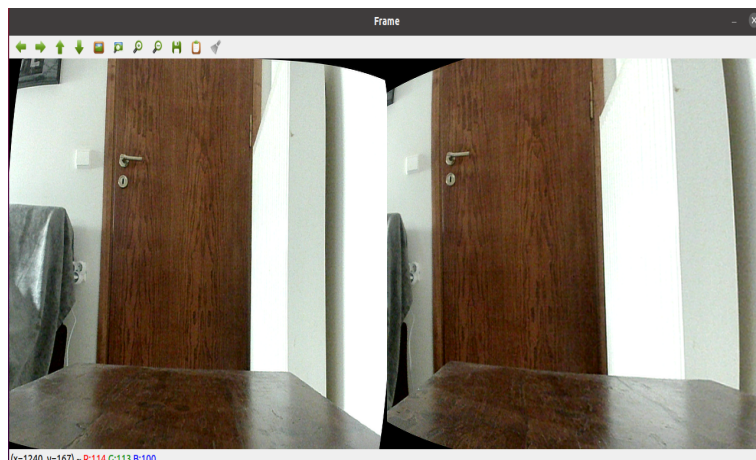
Για την υλοποίηση της στερεοσκοπικής ανόρθωσης στην εφαρμογή, έχουν χρησιμοποιηθεί οι συναρτήσεις που περιγράφηκαν εκτενώς στην προηγούμενη παράγραφο. Μέσω της εφαρμογής ο χρήστης μπορεί να εφαρμόσει την στερεοσκοπική ανόρθωση πολύ εύκολα στις εικόνες που προέρχονται από την στερεοσκοπική διάταξη.

Αρχικά ο χρήστης πρέπει να έχει εκτελέσει επιτυχώς, με αρκετά καλά αποτελέσματα, τα βήματα της βαθμονόμησης για το στερεοσκοπικού συστήματος, καθώς απαιτείται για την εκτέλεση της στερεοσκοπικής ανόρθωσης. Τα βήματα βαθμονόμησης περιγράφονται στο προηγούμενο κεφάλαιο. Οι παράμετροι που έχουν υπολογιστεί κατά τη διαδικασία της βαθμονόμησης αποθηκεύονται σε ένα αρχείο και στη συνέχεια χρησιμοποιούνται ως είσοδοι στις συναρτήσεις για τη στερεοσκοπική διόρθωση.



Σχήμα 39: Μενού Εφαρμογής, στερεοσκοπική ανόρθωση

Στη συνέχεια ο χρήστης πατώντας το κουμπί remap ξεκινάει τη διαδικασία της στερεοσκοπικής ανόρθωσης. Εμφανίζονται δυο παράθυρα που απεικονίζουν τις διορθωμένες εικόνες. Στο παρακάτω σχήμα αποικίζονται τα αποτελέσματα της στερεοσκοπικής ανόρθωσης.



Σχήμα 40: Αποτελέσματα της στερεοσκοπικής ανόρθωσης μέσω της εφαρμογής

Ο χρήστης πατώντας το κουμπί 'q' μπορεί να σταματήσει την εκτέλεση της διαδικασίας και να επιστρέψει στο αρχικό μενού της εφαρμογής.

### 3.7 Υλοποίηση στερεοσκοπικής συσχέτισης

Για την υλοποίηση της στερεοσκοπικής συσχέτισης χρησιμοποιήθηκαν οι αλγόριθμοι που περιγράφονται στο αντίστοιχο κεφάλαιο. Πιο συγκεκριμένα, οι αλγόριθμοι που χρησιμοποιήθηκαν για την υλοποίηση της εφαρμογής είναι:

- Stereo Block Matching
- Stereo Semi Global Block Matching

Η βιβλιοθήκη OpenCV προσφέρει υλοποιήσεις των αλγορίθμων αυτών. Γενικά, ανάλογα με τις απαιτήσεις και τις δυνατότητες του συστήματος επιλέγεται ο κατάλληλος αλγόριθμος. Επίσης, για τους αλγορίθμους αυτούς υπάρχουν μερικές παράμετροι που πρέπει να ρυθμιστούν κατάλληλα με βάση τις συνθήκες και τα αποτελέσματα που θέλουμε να εξάγουμε. Επιπλέον, η βιβλιοθήκη OpenCV παρέχει υλοποιήσεις των αλγορίθμων τόσο σε *CPU*, όσο και σε *Cuda*.

Στα πλαίσια της εφαρμογής αναπτύχθηκαν και υλοποιήθηκαν οι παραπάνω αλγόριθμοι σε *CPU* και *Cuda*. Έτσι, μέσω της εφαρμογής δίνεται η δυνατότητα στον χρήστη να τους συγκρίνει και να αποφανθεί ποιος είναι ο καταλληλότερος για την εφαρμογή του. Επίσης, ο χρήστης μπορεί να ρυθμίσει τις παραμέτρους των αλγορίθμων για τα μέγιστα δυνατά αποτελέσματα.

Είναι σημαντικό να αναφερθεί ότι πρέπει να έχουν εκτελεστεί με επιτυχία όλες οι απαραίτητες διαδικασίες, όπως η στερεοσκοπική βαθμονόμηση και η ανόρθωση των εικόνων, για καλύτερα αποτελέσματα και οι αλγόριθμοι να είναι πιο αποδοτικοί. Καθώς οι αλγόριθμοι ψάχνουν συγκεκριμένες περιοχές για τις αντιστοιχίες των σημείων, με αυτόν το τρόπο μειώνεται αρκετά ο χρόνος και η υπολογιστική ισχύς που χρειάζεται.

Ο χρήστης πατώντας το κουμπί 'Disparity Map' από το αρχικό μενού της εφαρμογής Σχήμα 39, θα του εμφανιστεί ένα νέο παράθυρο, μέσω του οποίου θα μπορεί να επιλέξει ποιον αλγόριθμο επιθυμεί να εφαρμόσει, καθώς επίσης να ρυθμίσει και τις παραμέτρους του.

Import Images

Right

Left

Show

3D Image

Video

Play Video

Stop Video

Stereo Matching Algorithms

☒ StereoBM

☐ StereoSGBM

☒ CPU

☐ CUDA

Apply

minDisparity	-100	<div></div>	100	<div>0</div>
numOfDisparities	16	<div></div>	256	<div>0</div>
BlockSize	5	<div></div>	255	<div>0</div>
speckleRange	0	<div></div>	40	<div>0</div>
speckleWindowSize	0	<div></div>	300	<div>0</div>
Disp12Max	0	<div></div>	0	<div>0</div>
uniquenessRatio	1	<div></div>	20	<div>0</div>
textureThreshold	1	<div></div>	1000	<div>0</div>
preFilterCap	1	<div></div>	63	<div>0</div>
PreFilterSize	5	<div></div>	255	<div>0</div>
P1	0	<div></div>	4000	<div>0</div>
P2	0	<div></div>	4000	<div>0</div>

Exit

Σχήμα 41: Εφαρμογή: Disparity map

Ο χρήστης έχει τις εξής δυνατότητες:

- Να εισάγει στερεοσκοπικές εικόνες,
- Να χρησιμοποιεί απευθείας το στερεοσκοπικό σύστημα και να εμφανίζονται τα αποτελέσματα σε πραγματικό χρόνο,
- Να επιλέγει τον αλγόριθμο που επιθυμεί να εφαρμοστεί.
- Να εξάγει ένα αρχείο το οποίο περιέχει όλα τα σημεία για την τρισδιάστατη ανακατασκευή της εικόνας.

### **Εισαγωγή στερεοσκοπικών εικόνων**

Η εισαγωγή των εικόνων πραγματοποιείται πατώντας το κουμπί Right για την εισαγωγή της δεξιάς εικόνας και αντίστοιχα πατώντας το κουμπί Left για την αριστερή εικόνα. Στη συνέχεια εμφανίζεται ένα παράθυρο με το οποίο ο χρήστης επιλέγει τις εικόνες που επιθυμεί. Πατώντας το κουμπί Show εμφανίζονται οι εικόνες οι οποίες έχουν υποστεί στερεοσκοπική ανόρθωση.

### **Χρήση στερεοσκοπικής κάμερας σε πραγματικό χρόνο**

Ο χρήστης πατώντας το κουμπί Play Video έχει τη δυνατότητα απευθείας να ανοίξει τις κάμερες του στερεοσκοπικού συστήματος και να εφαρμόζει σε πραγματικό χρόνο τους αλγορίθμους στερεοσκοπικής συσχέτισης και τις παραμέτρους τους. Πατώντας το κουμπί Stop Video τερματίζεται την λειτουργία αυτή.

### **Παράμετροι στερεοσκοπικών αλγορίθμων**

Η ρύθμιση των παραμέτρων των στερεοσκοπικών αλγορίθμων έχει καθοριστικό ρόλο στα αποτελέσματα. Η ρύθμιση δεν αποτελεί μια εύκολη διαδικασία, καθώς δεν υπάρχει κάποιος προκαθορισμένος τρόπος να πραγματοποιηθεί. Οι παράμετροι επηρεάζονται από τις συνθήκες που επικρατούν στη σκηνή και την ακρίβεια των αποτελεσμάτων που επιθυμούμε.

Μέσω της εφαρμογής ο χρήστης μπορεί να ρυθμίσει εύκολα τις παραμέτρους του αλγορίθμου που επιθυμεί σε πραγματικό χρόνο, καθώς οι ρυθμίσεις εφαρμόζονται απευθείας στην εικόνα.

Παρακάτω θα αναλυθούν εκτενέστερα οι παράμετροι των αλγορίθμων.

- **MinDisparity:** Η ελάχιστη τιμή που μπορεί να έχει η στερεοσκοπική διάφορα.
- **numOfDisparities:** Ο αριθμός των εικονοστοιχείων στα οποία γίνεται η αναζήτηση της αντιστοιχίας. Η τιμή πρέπει να είναι πάντα μεγαλύτερη του μηδενός ενώ πρέπει να διαιρείται με το 16. Όσο μεγαλύτερο είναι, τόσο αυξάνεται το υπολογιστικό κόστος.
- **BlockSize:** Το μέγεθος του παράθυρου για την αντιστοίχιση. Πρέπει να είναι περιττός αριθμός γιατί είναι συμμετρικό. Όσο μικρότερο είναι το παράθυρο, τόσο πιο λεπτομερής γίνεται η σκηνή.
- **Disp12Max:** Η μεγίστη επιτρεπομένη διάφορα μεταξύ των εικόνων
- **uniquenessRatio:** Παράμετρος που εφαρμόζεται στο post-filtering stage. Εάν η τιμή της στερεοσκοπικής διαφοράς δεν είναι αρκετά καλή από όλες τις άλλες διαφορές της περιοχές, τότε φιλτράρεται.
- **textureThreshold:** Το κατώτατο όριο της υψής, δηλαδή αποκλείονται περιοχές όπου δεν υπάρχει αρκετή υψή, ώστε να χαρακτηριστούν αξιόπιστες.
- **preFilterCap & preFilterSize:** Παράμετροι που εφαρμόζονται στο φίλτρο για την προετοιμασία της εικόνων (pre-filtering stage). Πιο συγκεκριμένα, το φίλτρο εξομαλύνει της φωτεινότητα και ενισχύει τις υψές των εικόνων για να υπάρχει καλύτερη συσχέτιση.
- **P1 & P2:** Παράμετροι ομαλοποίησης των στερεοσκοπικών διαφόρων

## 4 Περιγραφή εξοπλισμού

### 4.1 Περιγραφή ενσωματωμένου συστήματος

Το ενσωματωμένο σύστημα που προτείνεται να χρησιμοποιηθεί στην παρούσα διπλωματική εργασία είναι το Jetson Xavier NX της εταιρίας Nvidia. Κατά κύριο λόγο, το ενσωματωμένο σύστημα χρησιμοποιείται σε απαιτητικές εφαρμογές στο πεδίο της τεχνίτης νοημοσύνης και της μηχανικής όρασης, καθώς προσφέρει καλές επιδόσεις με τη βοήθεια των GPUs.



Σχήμα 42: Ενσωματωμένο σύστημα: Jetson Xavier NX. [11]

Τα κυρία χαρακτηριστικά του είναι:

- **GPU:** NVIDIA Volta architecture with 384 NVIDIA CUDA cores and 48 Tensor cores
- **CPU:** 6-core NVIDIA Carmel ARMv8.2 64-bit CPU 6 MB L2 + 4 MB L3
- **DL Accelerator:** 2x NVDLA Engines
- **Memory:** 8 GB 128-bit LPDDR4x 51.2GB/s
- **USB:** 4x USB 3.1, USB 2.0 Micro-B
- **Display:** HDMI and display port

## 4.2 Προδιαγραφές καμερών

Οι δύο κάμερες που χρησιμοποιούνται στη στερεοσκοπική διάταξη είναι απόλυτα όμοιες μεταξύ τους. Είναι τύπου Full HD Wide-angle WebCam. Τα βασικά χαρακτηριστικά τους είναι:

- Ανάλυση εικόνας: 1080p/720p
- FPS: 30
- Wide Viewing Angle: 77°



Σχήμα 43: Εικόνα κάμερας

### 4.3 Υλοποίηση στερεοσκοπικού συστήματος

Η στερεοσκοπική διάταξη που κατασκευάστηκε για τον σκοπό της εργασίας αυτής, ικανοποιεί την πιο απλή στερεοσκοπική γεωμετρία, η οποία βασίζεται σε δύο πανομοιότυπες κάμερες.

Για την επιλογή της απόστασης πραγματοποιήθηκαν αρκετές δοκιμές σύμφωνα με το άρθρο [1]. Οι συγγραφείς στην εργασία τους παρουσιάζουν ένα σύστημα το οποίο μετακινεί τις κάμερες ανάλογα με την απόσταση που ορίζεται και αναλύεται όλη η διαδικασία και ο σκοπός επιλογής της κατάλληλης απόστασης. Η απόσταση που επιλέχθηκε είναι 10cm και μπορεί να δώσει σχετικά ικανοποιητικά αποτελέσματα σε αντικείμενα που βρίσκονται σε απόσταση 50cm από το στερεοσκοπικό σύστημα. Οι κάμερες τοποθετήθηκαν και κολληθήκαν πάνω σε ένα κομμάτι ξύλο, ώστε να είναι δύσκολο να χαλάσει η γεωμετρία τους. Στο παρακάτω σχήμα φαίνεται η στερεοσκοπική διάταξη που υλοποιήθηκε.



Σχήμα 44: Στερεοσκοπική διάταξη καμερών

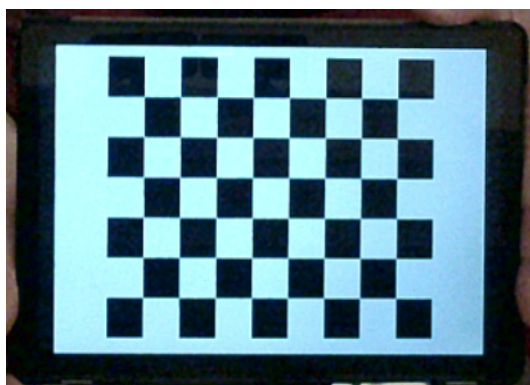
## 5 Αποτέλεσμα

Στα προηγούμενα κεφάλαια αναλύεται η θεωρία που χρειάζεται, όσον αφορά τη βαθμονόμηση του στερεοσκοπικού συστήματος, την ανόρθωση των εικόνων και τη στερεοσκοπική συσχέτιση. Επίσης, περιγράφεται η εφαρμογή και ο εξοπλισμός που αναπτύχθηκε για την τρισδιάστατη ανακατασκευή των εικόνων.

Σε αυτό το κεφάλαιο, θα αναλυθούν τα αποτελέσματα που προκύπτουν από την εφαρμογή σε συνδυασμό με το στερεοσκοπικό σύστημα. Στη συνέχεια, θα πραγματοποιηθεί σύγκριση μεταξύ των αλγορίθμων στερεοσκοπικής συσχέτισης. Τέλος, θα γίνει σύγκριση ανάμεσα στο ενσωματωμένο σύστημα που προτείνεται με έναν συμβατικό υπολογιστή.

### 5.1 Αποτελέσματα βαθμονόμησης στερεοσκοπικού συστήματος

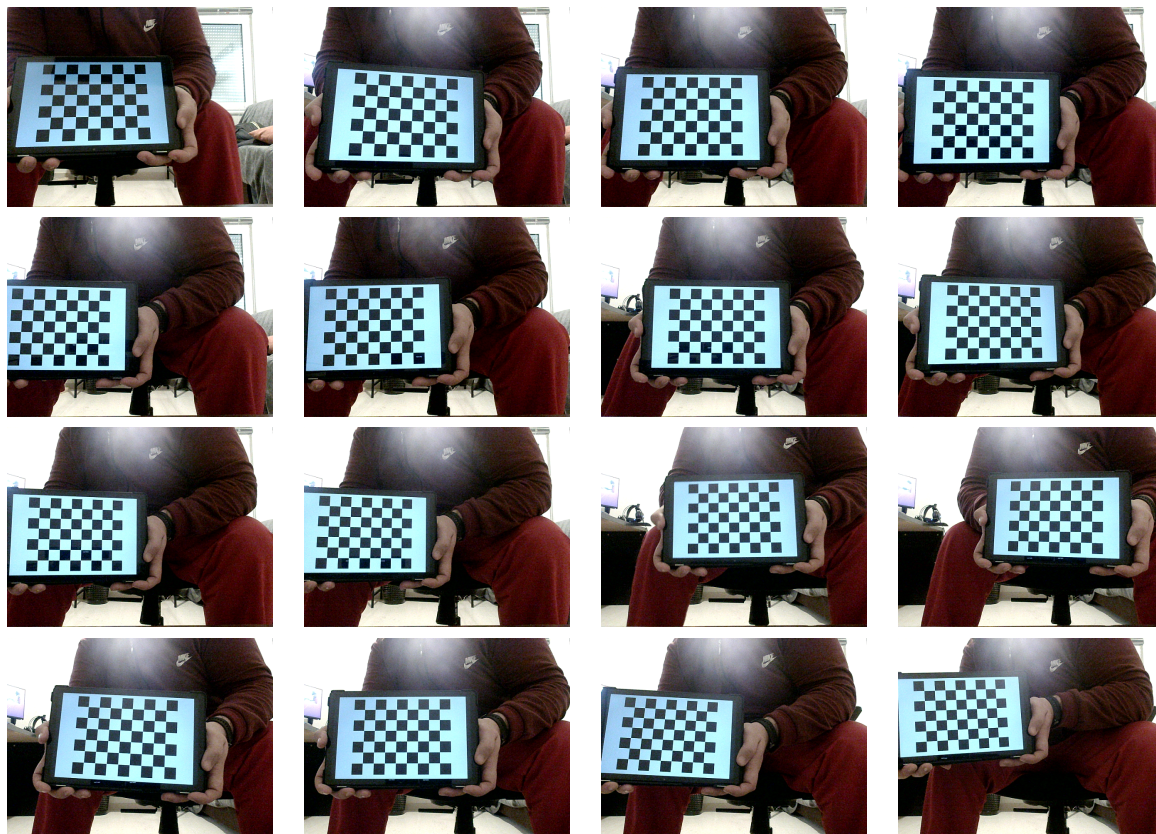
Όπως έχει ήδη αναφερθεί στη θεωρία σχετικά με τη βαθμονόμηση, πρέπει να κατασκευαστεί ένα αντικείμενο παρατήρησης, το οποίο πρέπει να είναι με ακρίβεια ορισμένο. Το αντικείμενο παρατήρησης που επιλέχθηκε είναι δισδιάστατο και τύπου σκακιέρας, όπως φαίνεται στο Σχήμα 15. Για την κατασκευή του αντικείμενου βαθμονόμησης χρησιμοποιήθηκαν εικόνες που απεικονίζουν το μοτίβο της σκακιέρας. Είναι εύκολο να βρεθούν τέτοιες εικόνες στο διαδίκτυο. Για την απεικόνιση του αντικείμενου βαθμονόμησης χρησιμοποιήθηκε ένα tablet, διότι είναι πιο εύκολο να δοκιμαστούν διάφορα μοτίβα, καθώς και πιο καλά ορισμένο σε σχέση με την εκτύπωση σε χαρτί. Το μοτίβο σκακιέρας που επιλέχθηκε είναι  $6 \times 8$  και το μέγεθος των τετράγωνων είναι  $20mm$ .



Σχήμα 45: Αντικείμενο βαθμονόμησης τύπου σκακιέρας

Στη συνέχεια, ακολουθήθηκε η απαραίτητη διαδικασία, όπως αυτή περιγράφεται παραπάνω. Πιο συγκεκριμένα, ορίστηκαν οι παράμετροι στην εφαρμογή που φορούν τα χαρακτηριστικά του μοτίβου, καθώς και πραγματοποιήθηκαν διαφορικές λήψεις του αντικείμενου από διάφορα σημεία. Η διαδικασία επαναλήφθηκε αρκετές φορές, ώστε να

προκύβουν τα καλύτερα αποτελέσματα. Έγιναν περίπου 50 λήψεις του αντικείμενου βαθμονόμησης.



Σχήμα 46: Ενδεικτικές λήψεις του αντικειμένου βαθμονόμησης

Τα αποτελέσματα της βαθμονόμησης του στερεοσκοπικού συστήματος φαίνονται παρακάτω.

Τα αποτελέσματα για την **αριστερή κάμερα** είναι:

$$\text{camera matrix} = \begin{bmatrix} 585.82 & 0 & 304.79 \\ 0 & 586.198 & 247.69 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\text{distortion coefficients} = [-0.0503 \quad 0.1220 \quad 1.035 \times 10^{-3} \quad 4.702 \times 10^{-4} \quad -0.1330]$$

Τα αποτελέσματα για τη **δεξιά κάμερα** είναι:

$$\text{camera matrix} = \begin{bmatrix} 604.87 & 0 & 303.27 \\ 0 & 605.86 & 247.69 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\text{distortion coefficients} = [-0.0361 \quad 0.1825 \quad 1.924 \times 10^{-3} \quad -1.413 \times 10^{-3} \quad -0.4273]$$

Τα αποτελέσματα για τη **στερεοσκοπική κάμερα** είναι:

Πίνακας μεταφοράς:

$$T = [119.26 \quad 3.717 \quad 8.993]$$

Πίνακας περιστροφής:

$$R = \begin{bmatrix} 0.999 & 0.0126 & -4.273 \times 10^{-3} \\ -0.0128 & 0.994 & -0.0296 \\ 3.894 \times 10^{-3} & 0.0297 & 0.9955 \end{bmatrix}$$

Το **σφάλμα (reprojection error)** που προκύπτει είναι:

Αριστερή κάμερα	0.135
Δεξιά κάμερα	0.145
Στερεοσκοπική κάμερα	0.761

## 5.2 Αποτελέσματα ανόρθωσης εικόνων

Παρακάτω παρουσιάζονται τα αποτελέσματα της στερεοσκοπικής ανόρθωσης.



Σχήμα 47: Αρχικές εικόνες από το στερεοσκοπικό σύστημα



Σχήμα 48: Εικόνες από το στερεοσκοπικό σύστημα μετά την εφαρμογή της στερεοσκοπικής ανόρθωσης

### 5.3 Αποτελέσματα αλγορίθμων στερεοσκοπικής συσχέτισης

Καθοριστικό ρόλο στα αποτελέσματα της στερεοσκοπικής συσχέτισης παίζει η στερεοσκοπική ανόρθωση. Επιπλέον, συμβάλει άμεσα στη μείωση της υπολογιστής ισχύος που απαιτείται από τους αλγορίθμους για την εύρεση και την αντιστοιχία των σημείων. Επίσης, πρέπει να αναφερθεί ότι τα αποτελέσματα διαμορφώνονται ανάλογα με τις ρυθμίσεις των παραμέτρων που επιλέγονται, μέσω της εφαρμογής.

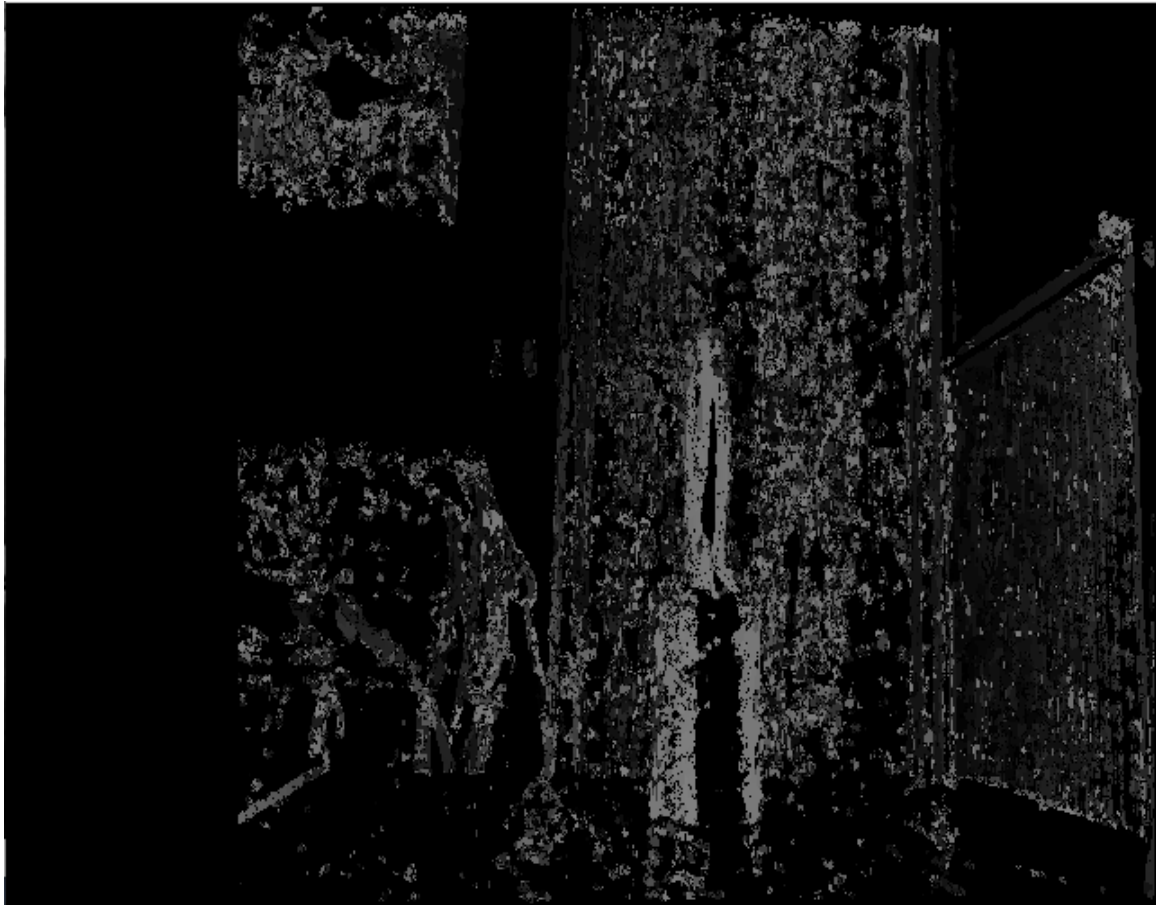
Παρακάτω παρουσιάζονται τα αποτελέσματα των αλγορίθμων που προέκυψαν από το στερεοσκοπικό σύστημα σε συνδυασμό με το ενσωματωμένο σύστημα. Αρχικά, θα αναλυθούν τα αποτελέσματα του Stereo Block Matching τόσο σε CPU, όσο και σε CUDA. Έπειτα, θα παρουσιαστούν τα αποτελέσματα του Semi Global Block Matching. Ο χρήστης έχει τη δυνατότητα να καθορίσει το disparity μέσω του γραφικού περιβάλλοντος που προσφέρεται από την εφαρμογή, αυξάνοντας το υπολογιστικό κόστος. Για το λόγο αυτό, μετά από αρκετές δοκιμές προτείνεται η χρήση της τιμής 128. Πρέπει να αναφερθεί ότι η τρισδιάστατη απεικόνιση της σκηνής γίνεται με τη χρήση του προγράμματος Meshlab. Πιο συγκεκριμένα, φορτώνεται το αρχείο με τα 3D σημεία της σκηνής, το οποίο έχει παραχθεί από την εφαρμογή. Τέλος, πραγματοποιείται σύγκριση μεταξύ του ενσωματωμένου και ενός υπολογιστή.

#### Stereo Block Matching (CPU)

Παρακάτω απεικονίζονται τα αποτελέσματα που προέκυψαν από τον αλγόριθμο Stereo Block Matching με τη χρήση CPU. Στο Σχήμα 49 παρουσιάζεται πως ρυθμίστηκαν οι παράμετροι του αλγορίθμου μέσω της εφαρμογής. Στο Σχήμα 50 απεικονίζεται το disparity map που δημιουργήθηκε από τις εικόνες του Σχήματος 48. Παρατηρείται ότι το αντικείμενο που είναι πιο κοντά στην κάμερα έχει πιο έντονη απόχρωση γκρι και είναι πιο φωτεινό από τα άλλα αντικείμενα που βρίσκονται πιο μακριά.

minDisparity	-100	<input type="range"/>	100	0
numOfDisparities	16	<input type="range"/>	256	128
BlockSize	5	<input type="range"/>	255	5
speckleRange	0	<input type="range"/>	40	22
speckleWindowSize	0	<input type="range"/>	300	0
Disp12Max	0	<input type="range"/>	0	6400
uniquenessRatio	1	<input type="range"/>	20	4
textureThreshold	1	<input type="range"/>	1000	616
preFilterCap	1	<input type="range"/>	63	63
PreFilterSize	5	<input type="range"/>	255	255
P1	0	<input type="range"/>	4000	0
P2	0	<input type="range"/>	4000	0

Σχήμα 49: Διαμόρφωση παραμέτρων του αλγορίθμου Stereo Block Matching (*CPU*)



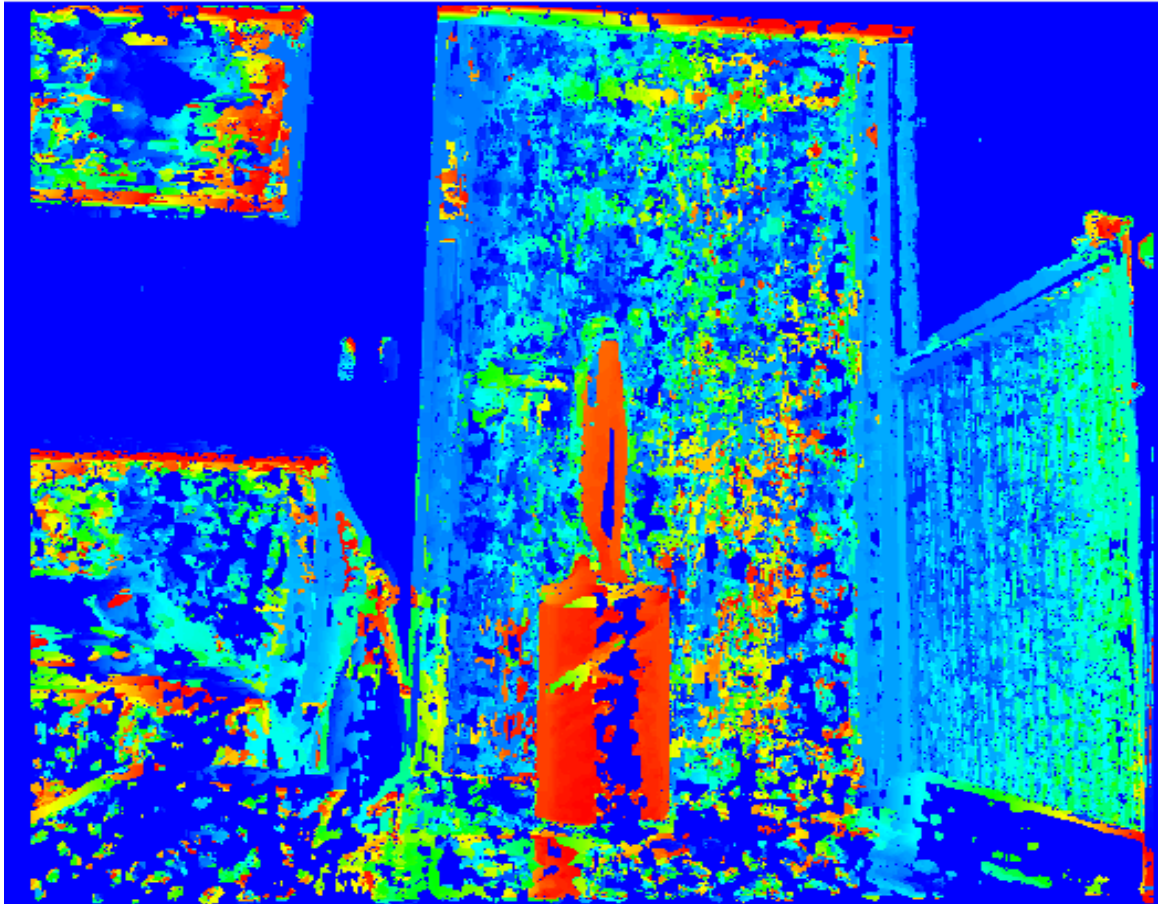
Σχήμα 50: Disparity map αλγορίθμου Stereo Block Matching (*CPU*)

## Stereo Block Matching (CUDA)

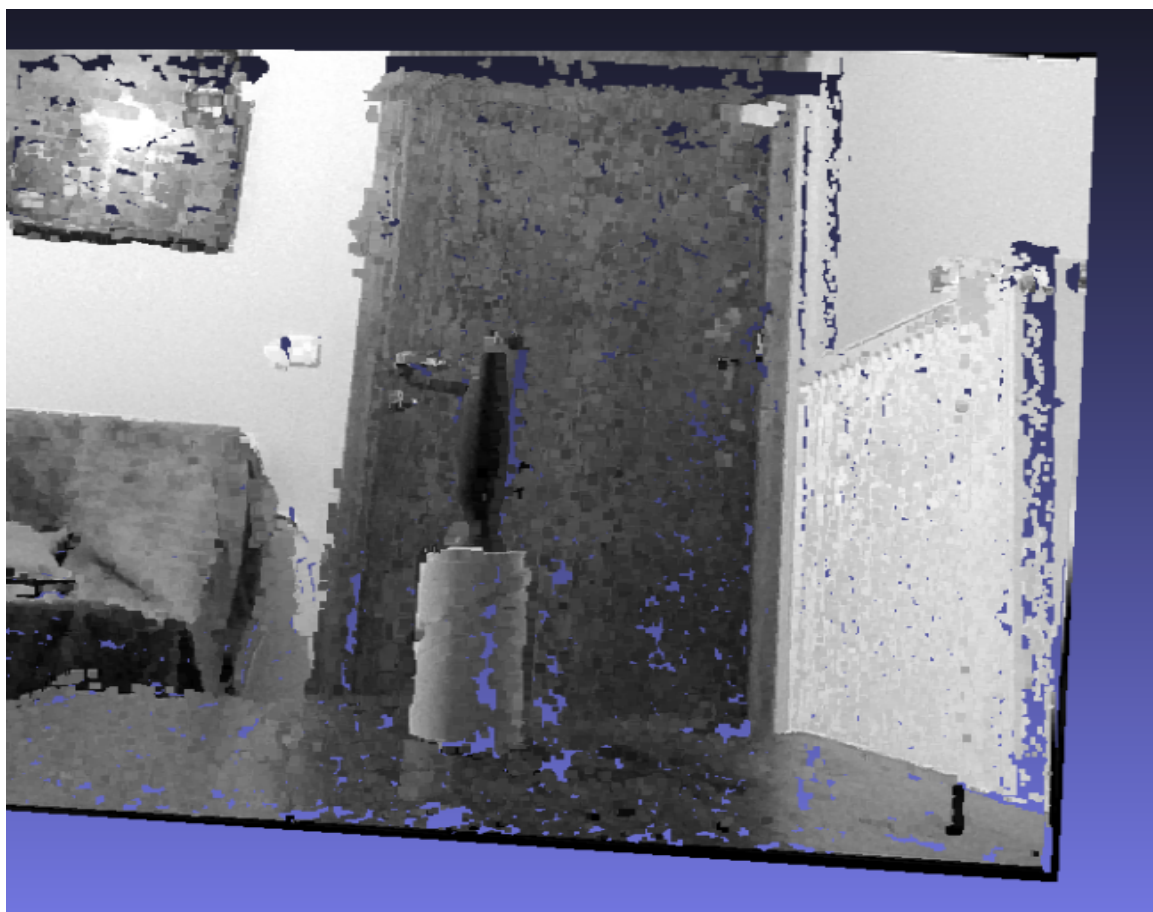
Παρακάτω απεικονίζονται τα αποτελέσματα που προέκυψαν από τον αλγόριθμο Stereo Block Matching με τη χρήση *GPUs*. Επίσης, στο Σχήμα 51 παρουσιάζεται πως ρυθμίστηκαν οι παράμετροι του αλγορίθμου μέσω της εφαρμογής. Στο Σχήμα 52 απεικονίζεται το disparity map που δημιουργήθηκε από τις εικόνες του Σχήματος 48. Παρατηρείται ότι το αντικείμενο που έχει κόκκινο χρώμα βρίσκεται κοντά στην κάμερα, ενώ τα αντικείμενα που έχουν πράσινο χρώμα βρίσκονται λίγο πιο πίσω από το κόκκινο αντικείμενο, τα αντικείμενα που έχουν μπλε χρώμα βρίσκονται μακριά από την κάμερα. Οι αστοχίες που διακρίνονται οφείλονται στις δυσκολίες που αντιμετωπίζει ο αλγόριθμος λόγω των διαφορετικών υφών της σκηνής. Το αντικείμενο που βρίσκεται στα δεξιά της σκηνής (καλοριφέρ), είναι σχετικά ομοιόμορφο και η απεικόνισή του είναι ικανοποιητική. Το αντικείμενο που βρίσκεται πίσω από το αντικείμενο που απεικονίζεται με κόκκινο χρώμα (πόρτα), λόγω των διαφορετικών υφών που παρουσιάζει έχει και τις περισσότερες αστοχίες. Το χρώμα που επικρατεί είναι το μπλε, άρα υπάρχει η αίσθηση του βάθους. Στο Σχήμα 53 απεικονίζεται η τρισδιάστατη ανακατασκευή της σκηνής.

minDisparity	-100		100	0
numOfDisparities	16		256	128
BlockSize	5		255	5
speckleRange	0		40	0
speckleWindowSize	0		300	0
Disp12Max	0		0	0
uniquenessRatio	1		20	1
textureThreshold	1		1000	14
preFilterCap	1		63	1
PreFilterSize	5		255	7
P1	0		4000	0
P2	0		4000	0

Σχήμα 51: Διαμόρφωση παραμέτρων του αλγορίθμου Stereo Block Matching (*CUDA*)





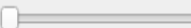
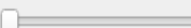
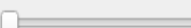
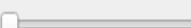
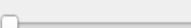
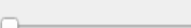
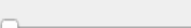
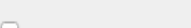
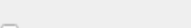
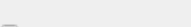
Σχήμα 52: Disparity map αλγορίθμου Stereo Block Matching (*CUDA*)



Σχήμα 53: Point Cloud αλγορίθμου Stereo Block Matching (*CUDA*) μέσω MeshLab

## Semi Global Stereo Matching

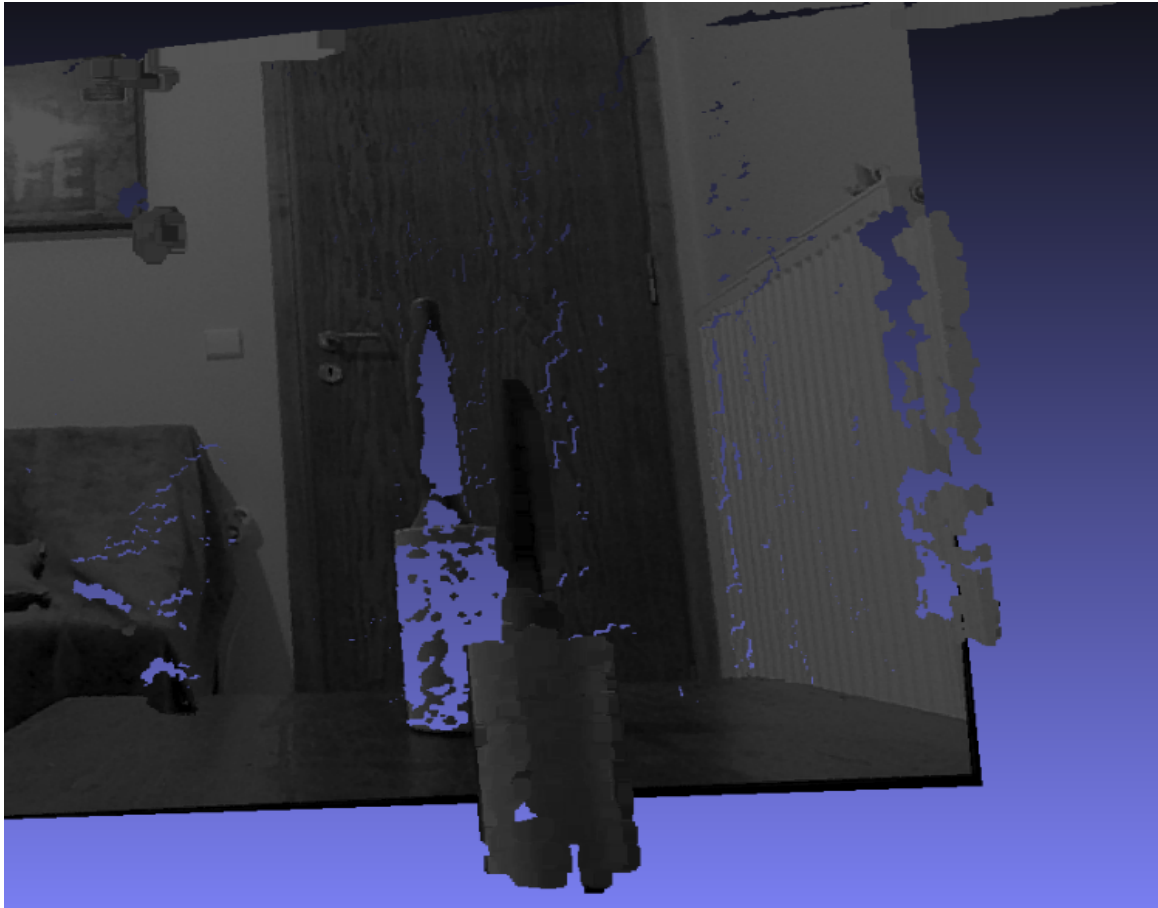
Παρακάτω απεικονίζονται τα αποτελέσματα που προέκυψαν από τον αλγόριθμο Semi Global Stereo Matching. Στο Σχήμα 54 παρουσιάζεται πως ρυθμίστηκαν οι παράμετροι του αλγορίθμου μέσω της εφαρμογής. Στο Σχήμα 55 απεικονίζεται το disparity map που δημιουργήθηκε από τις εικόνες της σκηνής του Σχήματος 48. Παρατηρείται ότι το αντικείμενο που βρίσκεται σε κοντινή απόσταση από την κάμερα, έχει έντονο φωτεινό γκρι χρώμα, τα αντικείμενα που βρίσκονται πιο μακριά έχουν πιο σκούρο χρώμα. Στο Σχήμα 56 παρουσιάζεται η τρισδιάστατη ανακατασκευή της σκηνής μετά από κατάλληλο φιλτράρισμα.

minDisparity	-100		100	0
numOfDisparities	16		256	128
BlockSize	5		255	5
speckleRange	0		40	0
speckleWindowSize	0		300	0
Disp12Max	0		0	0
uniquenessRatio	1		20	0
textureThreshold	1		1000	0
preFilterCap	1		63	0
PreFilterSize	5		255	0
P1	0		4000	0
P2	0		4000	0

Σχήμα 54: Διαμόρφωση παραμέτρων του αλγορίθμου Semi Global Stereo Matching



Σχήμα 55: Disparity map αλγορίθμου Semi Global Stereo Matching



Σχήμα 56: Point Cloud αλγορίθμου Semi Global Stereo Matching μέσω MeshLab

### Χρόνος υπολογισμού του Disparity map

Ο χρόνος που χρειάζεται ο κάθε αλγόριθμος για τον υπολογισμό του Disparity map παρουσιάζεται παρακάτω. Ο πίνακας 1 περιέχει τα αποτελέσματα που προέκυψαν από το ενσωματωμένο σύστημα, ενώ ο πίνακας 2 περιέχει τα αποτελέσματα από τον υπολογιστή. Ο υπολογιστής έχει επεξεργαστή Intel Core i5-4440, μνήμη RAM 12GB DDR3 και κάρτα γραφικών NVIDIA GeForce GTX 1660.

Resolution	Stereo BM CPU	Stereo BM CUDA	Stereo SGBM
640p	0.065s	0.010s	0.354s
1080p	0.236s	0.056s	2.153s

Πίνακας 1: Χρόνος υπολογισμού του Disparity map για το ενσωματωμένο σύστημα

Resolution	Stereo BM CPU	Stereo BM CUDA	Stereo SGBM
640p	0.015s	0.002s	0.158s
1080p	0.060s	0.012s	1.011s

Πίνακας 2: Χρόνος υπολογισμού του Disparity map για τον υπολογιστή

Παρατηρείται ότι ο αλγόριθμος Semi Global Block Matching για τον υπολογισμό του disparity map στο ενσωματωμένο σύστημα χρειάζεται περίπου 0.3s για εικόνες ανάλυσης 640p και 2.1s για 1080p. Για τον υπολογιστή χρειάζεται περίπου 0.1s για 640p και 1.4s για 1080p. Επίσης, ο αλγόριθμος Stereo Block Matching σε CPU για το υπολογισμό του disparity map στο ενσωματωμένο σύστημα χρειάζεται περίπου 0.05s για εικόνες ανάλυσης 640p και 0.2s για 1080p. Για τον υπολογιστή χρειάζεται περίπου 0.015s για 640p και 0.060s για 1080p. Ο αλγόριθμος Stereo Block Matching σε CUDA για τον υπολογισμό του disparity map στο ενσωματωμένο σύστημα χρειάζεται περίπου 0.01s για εικόνες ανάλυσης 640p και 0.05s για 1080p. Για τον υπολογιστή χρειάζεται περίπου 0.002s για 640p και 0.01s για τα 1080p.

## 6 Συμπεράσματα και προτάσεις για μελλοντική έρευνα

### Συμπεράσματα

Σκοπό της διπλωματικής εργασίας αποτέλεσε η ανάπτυξη και υλοποίηση μιας ολοκληρωμένης εφαρμογής για την τρισδιάστατη ανακατασκευή εικόνων σε ενσωματωμένο σύστημα, αξιοποιώντας κάμερες χαμηλού κόστους. Επίσης, το γραφικό περιβάλλον που αναπτύχθηκε επιτρέπει στο χρήστη να εκτελέσει εύκολα όλες τις διαδικασίες που απαιτούνται, όπως η βαθμονόμηση του στερεοσκοπικού συστήματος και η στερεοσκοπική ανόρθωση των εικόνων. Επιπλέον, προσφέρεται η δυνατότητα επιλογής αλγόριθμου για τη δημιουργία του χάρτη βάθους της σκηνής που ικανοποιεί τις απαιτήσεις του χρήστη. Επιπρόσθετα, αποτελεί ιδιαίτερο χαρακτηριστικό της εφαρμογής το γεγονός ότι οι μεταβολές που επιφέρει η ρύθμιση των παραμέτρων του αλγορίθμου εφαρμόζονται σε πραγματικό χρόνο. Τέλος, η εφαρμογή αναπτύχθηκε με τέτοιο τρόπο ώστε να αξιοποιεί αποτελεσματικά τους πόρους που προσφέρει το εκάστοτε σύστημα, ενώ η υλοποίηση της εφαρμογής έγινε σε γλώσσα C++.

Τα αποτελέσματα που προέκυψαν κατά τη διεξαγωγή των δοκίμων τόσο στο ενσωματωμένο σύστημα, όσο και στον υπολογιστή οδήγησαν στα ακόλουθα συμπεράσματα:

- Ο αλγόριθμος Semi Global Block Matching για τον υπολογισμό του disparity map χρειάζεται αρκετό χρόνο και σχετικά μεγάλη υπολογιστή ισχύ. Σημαντικό πλεονέκτημα του αλγορίθμου αυτού είναι τα πολύ καλά αποτελέσματα που προσφέρει. Συνεπώς, η χρήση του προτείνεται σε περιβάλλοντα με σχετικά αργές μεταβολές.
- Ο αλγόριθμος Stereo Block Matching σε CPU για τον υπολογισμό του disparity map απαιτεί λιγότερο χρόνο και λιγότερη υπολογιστή ισχύ από τον Semi Global Block Matching. Συνεπώς, μπορεί να χρησιμοποιηθεί σε περιπτώσεις όπου οι μεταβολές της σκηνής είναι σχετικά γρήγορες, καθώς ο χρόνος υπολογισμού του χάρτη βάθους είναι μικρός. Επίσης, μπορεί να χρησιμοποιηθεί σε εφαρμογές όπου η υπολογιστική ισχύς είναι περιορισμένη και η τρισδιάστατη ανακατασκευή δεν αποτελεί πρωταρχικό στόχο. Για παράδειγμα, το στερεοσκοπικό σύστημα σε συνδυασμό με τον αλγόριθμο αυτό μπορεί να λειτουργήσει ως αισθητήρας βάθους, υπολογίζοντας σε πραγματικό χρόνο την απόσταση των αντικείμενων από την κάμερα.
- Τα βέλτιστα αποτελέσματα για τον υπολογισμό του disparity map επιτυγχάνονται με χρήση του αλγορίθμου Stereo Block Matching σε CUDA, καθώς σε μικρούς χρόνους επεξεργασίας αποδίδονται ικανοποιητικά αποτελέσματα. Λόγω της υπολογιστικής ισχύς που προσφέρουν τα GPUs, ο αλγόριθμος μπορεί να χρησιμοποιηθεί για την τρισδιάστατη ανακατασκευή της σκηνής σε πραγματικό χρόνο. Επίσης,

μπορεί να χρησιμοποιηθεί σε απαιτητικές εφαρμογές όπου οι μεταβολές του περιβάλλοντος είναι γρήγορες και απαιτείται ακρίβεια και λεπτομέρεια.

- Το ενσωματωμένο σύστημα Jetson Xavier που επιλέχθηκε είναι ένα από τα καλύτερα που διαθέτει η εταιρία Nvidia, καθώς προσφέρει μεγάλη υπολογιστή ισχύ. Ωστόσο, ο επεξεργαστής που διαθέτει δεν είναι τόσο αποδοτικός, καθώς δεν έχει κατασκευαστεί για να εκτελεί δύσκολες και απαιτητικές διεργασίες. Πάραυτα, μπορεί να χρησιμοποιηθεί για τη δημιουργία του χάρτη βάθους σε πραγματικό χρόνο με καλά αποτελέσματα. Οι αποδόσεις των GPUs του ενσωματωμένου συστήματος είναι πολύ ικανοποιητικές, καθώς προσφέρουν μεγάλη υπολογιστή ισχύ, καταναλώνοντας χαμηλή ενέργεια. Αντίθετα, οι αποδόσεις του υπολογιστή είναι καλύτερες από το ενσωματωμένο σύστημα, όμως απαιτείται μεγαλύτερη κατανάλωση ενέργειας.
- Η εργασία του Chand αναφέρεται στη διαδικασία της τρισδιάστατης ανακατασκευής της σκηνής με τη χρήση ενός στερεοσκοπικού συστήματος χρησιμοποιώντας βιβλιοθήκες ανοιχτού κωδικά. Παρόλο που περιγράφεται αναλυτικά η διαδικασία, είναι σχετικά περίπλοκη και απαιτεί καλές γνώσεις προγραμματισμού. Ωστόσο, η διαδικασία αυτή απλοποιήθηκε σημαντικά, με τη βοήθεια του γραφικού περιβάλλοντος που διαθέτει η εφαρμογή, η οποία αναπτύχθηκε στην παρούσα διπλωματική. Οι Ildar και Romazon προτείνουν ένα ολοκληρωμένο σύστημα χαμηλού κόστους για τον υπολογισμό του χάρτη βάθους της σκηνής. Επίσης, έχουν υλοποιήσει μια ολοκληρωμένη εφαρμογή, για την εύκολη ρύθμιση των παραμέτρων του αλγορίθμου. Περιορισμός της εφαρμογής αυτής είναι ότι διαθέτει έναν αλγόριθμο για τον υπολογισμό του χάρτη βάθους. Αντίθετα, στην εφαρμογή, που προτείνεται σε αυτή τη διπλωματική, προσφέρεται η δυνατότητα επιλογής αλγορίθμων ανάλογα με τις δυνατότητες του συστήματος και τα αποτελέσματα που επιθυμεί ο χρήστης. [6] [4]
- Συγκρίνοντας την απόδοση του συστήματος με άλλες τεχνολογίες, οι οποίες χρησιμοποιούνται στο πεδίο της μηχανικής όρασης για το υπολογισμό του χάρτη βάθους, διαπιστώνεται τα εξής Στο άρθρο [14] οι Απόστολος Δόλλας και Κυπριανός Παπαδημητρίου προτείνουν μια υλοποίηση του αλγορίθμου στερεοσκοπικής συσχέτισης σε τεχνολογία FPGA. Η απόδοση της υλοποίησης αυτής δίνει αρκετά καλά αποτελέσματα, για εικόνες υψηλής ανάλυσης, καθώς η απόδοση είναι στα 87fps. Αντίθετα, η απόδοση του συστήματος, το οποίο χρησιμοποιείται στα πλαίσια της διπλωματικής εργασίας για εικόνες υψηλής ανάλυσης, είναι στα 17fps. Συνεπώς, η χρήση GPUs δεν ενδείκνυνται για την επίλυση του προβλήματος.

Συμπερασματικά, η εφαρμογή που αναπτύχθηκε στα πλαίσια της παρούσας εργασίας επιτρέπει την ικανοποιητική τρισδιάστατη ανακατασκευή των αντικείμενων με τη χρήση χαμηλών κόστους, αξιοποιώντας τις δυνατότητες που προσφέρει ένα ενσωματωμένο σύστημα.

## Περιορισμοί και προτάσεις για μελλοντική έρευνα

Οι περιορισμοί και οι προτάσεις για μελλοντική εργασία συνοψίζονται ακολούθως:

- Βασικός περιορισμός αποτέλεσε η δυσκολία ενσωμάτωσης περισσότερων αλγορίθμων για την δημιουργία του χάρτη βάθους. Ωστόσο, θα ήταν χρήσιμο να ενσωματωθούν και άλλοι αλγόριθμοι προσφέροντας περισσότερες επιλογές, καθώς η εφαρμογή που υλοποιήθηκε στα πλαίσια της διπλωματικής είναι επεκτάσιμη.
- Όσον αφορά την ελαχιστοποίηση του οικονομικού κόστους, θα μπορούσε να χρησιμοποιηθεί ένα πιο φθηνό ενσωματωμένο σύστημα και να διερευνηθούν οι αποδόσεις του.
- Ιδιαίτερο ενδιαφέρον θα αποτελούσε η διεξαγωγή δοχίμων χρησιμοποιώντας κάμερες υψηλής ανάλυσης, ώστε να μελετηθεί η επίδραση στην ποιότητα των αποτελεσμάτων.
- Ήταν μια πρώτη υλοποίηση της για να δοκιμαστεί το περιβάλλον της εφαρμογής στο συγκεκριμένο ενσωματωμένο σύστημα, στο μέλλον μπορεί να εξελιχθεί με καλύτερες αρχιτεκτονικές.
- Τέλος, σημαντική βελτίωση της εφαρμογής θα αποτελούσε η δυνατότητα τρισδιάστατης απεικόνισης σε πραγματικό χρόνο.

## Βιβλιογραφία

- [1] Wutthigrai Boonsuk. Investigating effects of stereo baseline distance on accuracy of 3d projection for industrial robotic applications. In *5th IAJC/ISAM joint international conference*, page 9, 2016.
- [2] Boguslaw Cyganek and J Paul Siebert. *An introduction to 3D computer vision techniques and algorithms*. John Wiley & Sons, 2011.
- [3] STEVEN B GOLDBERG. Larry matthies stereo vision and rover navigation software for planetary exploration. In *IEEE Conference Proceedings, March 2002, Big Sky, Montana, USA*, 2002.
- [4] Mohan Chand Hanumara. Calibrating and creating point cloud from a stereo camera setup using opencv. [https://www.researchgate.net/publication/308325524\\_Calibrating\\_and\\_Creating\\_Point\\_Cloud\\_from\\_a\\_Stereo\\_Camera\\_Setup\\_Using\\_OpenCV](https://www.researchgate.net/publication/308325524_Calibrating_and_Creating_Point_Cloud_from_a_Stereo_Camera_Setup_Using_OpenCV), 2022 (accessed December 2021).
- [5] Heiko Hirschmuller. Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on pattern analysis and machine intelligence*, 30(2):328–341, 2007.
- [6] Rakhmatulin Ildar and Eugene Pomazov. Low-cost stereovision system (disparity map) for few dollars. *arXiv preprint arXiv:2106.00905*, 2021.
- [7] Χριστόφορος Κανελλάκης and Γεώργιος Κυρίτσης. *Ανάπτυξη διάταξης βιομηχανικής όρασης για προσδιορισμό θέσης και προσανατολισμού κινούμενων αντικειμένων και οδήγηση ρομποτικού βραχίονα*. PhD thesis, 2014.
- [8] Bogdan Kwolek and Michal Kepski. Human fall detection on embedded platform using depth maps and wireless accelerometer. *Computer methods and programs in biomedicine*, 117(3):489–501, 2014.
- [9] Chaohui Lü, Xi Wang, and Yinghua Shen. A stereo vision measurement system based on opencv. In *2013 6th International Congress on Image and Signal Processing (CISP)*, volume 2, pages 718–722. IEEE, 2013.
- [10] Chris Murphy, Daniel Lindquist, Ann Marie Rynning, Thomas Cecil, Sarah Leavitt, and Mark L Chang. Low-cost stereo vision on an fpga. In *15th Annual IEEE*

- Symposium on Field-Programmable Custom Computing Machines (FCCM 2007)*, pages 333–334. IEEE, 2007.
- [11] nvidia. Jetson xavier nx. <https://developer.nvidia.com/embedded/jetson-xavier-nx-devkit>, 2022 (accessed December 2021).
  - [12] Fabio Oleari, Dario Lodi Rizzini, and Stefano Caselli. A low-cost stereo system for 3d object recognition. In *2013 IEEE 9th international conference on intelligent computer communication and processing (ICCP)*, pages 127–132. IEEE, 2013.
  - [13] OpenCV. Camera calibration and 3d reconstruction. [https://docs.opencv.org/4.x/d9/db7/tutorial\\_py\\_table\\_of\\_contents\\_calib3d.html](https://docs.opencv.org/4.x/d9/db7/tutorial_py_table_of_contents_calib3d.html), 2022 (accessed December 2021).
  - [14] Kyprianos Papadimitriou, Sotiris Thomas, and Apostolos Dollas. An fpga-based real-time system for 3d stereo matching, combining absolute differences and census with aggregation and belief propagation. In *IFIP/IEEE International Conference on Very Large Scale Integration-System on a Chip*, pages 168–187. Springer, 2013.
  - [15] Σοφία Φάκα. *Τρισδιάστατη ανακατασκευή αντικειμένου από φωτογραφίες (με χρήση Matlab)*. PhD thesis, 2010.
  - [16] Roberto Pisapia. Disparity map extraction for a low cost 3d sensor. <https://members.loria.fr/SATabbone/Disparity%20map%20extraction%20for%20a%20low%20cost%203D%20sensor%20-%20Pisapia%20Roberto.pdf>, 2022 (accessed December 2021).
  - [17] G Pomaska. Stereo vision applying opencv and raspberry pi. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42:265–269, 2019.
  - [18] Satyarth Praveen. Sefficient depth estimation using sparse stereo-vision with other perception techniques. <https://www.intechopen.com/chapters/67181>, 2022 (accessed December 2021).
  - [19] Wikipedia. Computer vision. [https://el.wikipedia.org/wiki/Computer\\_Vision](https://el.wikipedia.org/wiki/Computer_Vision), 2022 (accessed December 2021).
  - [20] Wikipedia. Mars rover. [https://en.wikipedia.org/wiki/Mars\\_rover](https://en.wikipedia.org/wiki/Mars_rover), 2022 (accessed December 2021).

- [21] Wikipedia. Pinhole camera. [https://en.wikipedia.org/wiki/Pinhole\\_camera](https://en.wikipedia.org/wiki/Pinhole_camera), 2022 (accessed December 2021).
- [22] Jure Zbontar and Yann LeCun. Computing the stereo matching cost with a convolutional neural network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1592–1599, 2015.
- [23] Zhengyou Zhang. A flexible new technique for camera calibration. *IEEE Transactions on pattern analysis and machine intelligence*, 22(11):1330–1334, 2000.